



(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:  
**02.12.2020 Bulletin 2020/49**

(51) Int Cl.:  
**G10L 19/24<sup>(2013.01)</sup> G10L 21/038<sup>(2013.01)</sup>**

(21) Application number: **20185468.4**

(22) Date of filing: **15.09.2011**

(84) Designated Contracting States:  
**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**

(30) Priority: **15.09.2010 KR 20100090582**  
**22.10.2010 KR 20100103636**  
**29.12.2010 KR 20100138045**

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:  
**16172268.1 / 3 113 182**  
**11825447.3 / 2 617 033**

(71) Applicant: **Samsung Electronics Co., Ltd.**  
**Gyeonggi-do 16677 (KR)**

(72) Inventors:  
• **SUNG, Ho Sang**  
**Gyeonggi-do (KR)**  
• **CHOO, Ki Hyun**  
**Seoul (KR)**  
• **OH, Eun Mi**  
**Seoul (KR)**

(74) Representative: **Appleyard Lees IP LLP**  
**15 Clare Road**  
**Halifax HX1 2HY (GB)**

Remarks:

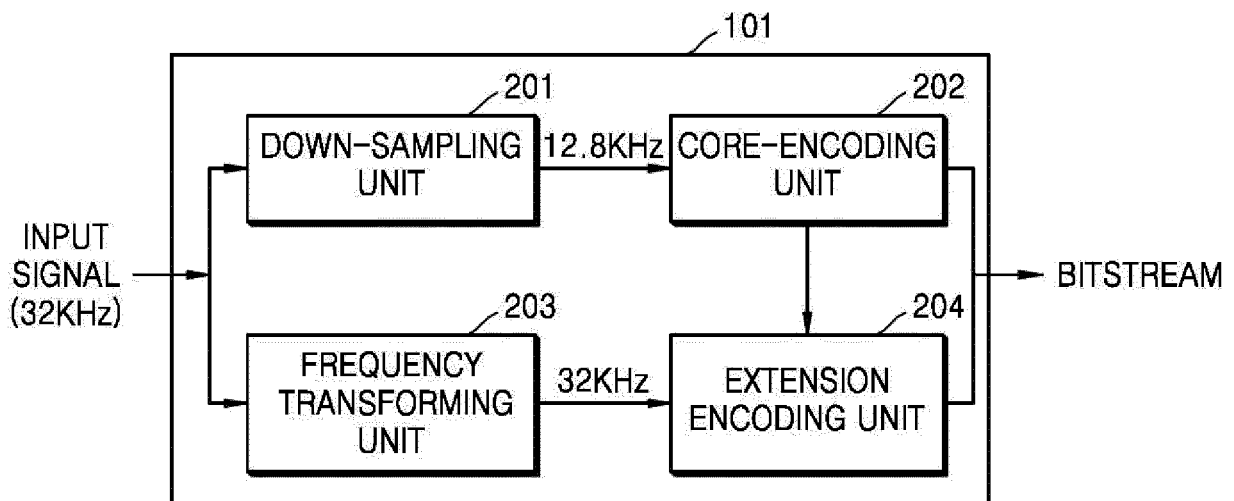
This application was filed on 13.07.2020 as a divisional application to the application mentioned under INID code 62.

(54) **APPARATUS AND METHOD FOR ENCODING SIGNAL FOR HIGH FREQUENCY BANDWIDTH EXTENSION**

(57) An apparatus and method for encoding and decoding a signal for high frequency bandwidth extension are provided. An encoding apparatus may down-sample a time domain input signal, may core-encode the down-sampled time domain input signal, may transform

the core-encoded time domain input signal to a frequency domain input signal, and may perform bandwidth extension encoding using a basic signal of the frequency domain input signal.

**FIG. 2**



**Description**

## Technical Field

- 5 **[0001]** One or more example embodiments of the following description relate to a method and apparatus for encoding or decoding an audio signal such as a speech signal or a music signal, and more particularly, to a method and apparatus for encoding or decoding a signal corresponding to a high-frequency domain among audio signals.

## Background Art

- 10 **[0002]** A signal corresponding to a high-frequency domain is less sensitive to a fine structure of a frequency than a signal corresponding to a low-frequency domain. Accordingly, there is a need to increase an encoding efficiency to overcome a restriction of bits available when encoding an audio signal. Thus, a large number of bits may be allocated to a signal corresponding to a low-frequency domain, while a smaller number of bits may be allocated to a signal  
15 corresponding to a high-frequency domain.

**[0003]** Such a scheme may be applied to a Spectral Band Replication (SBR) technology. SBR technology may be used to improve encoding efficiency by representing high-band component signals as an envelope, and by synthesizing the high-band component signals during the decoding of the high-band component signals, based on a fact that an auditory sense of a human being has a relatively low resolution in a high-band signal.

- 20 **[0004]** In SBR technology, there is a demand for an improved method for extending a bandwidth of a high-frequency domain.

## Disclosure of Invention

- 25 **[0005]** The foregoing and/or other aspects are achieved by providing an encoding apparatus including a down-sampling unit to down-sample a time domain input signal, a core-encoding unit to core-encode the down-sampled time domain input signal, a frequency transforming unit to transform the core-encoded time domain input signal to a frequency domain input signal, and an extension encoding unit to perform bandwidth extension encoding using a basic signal of the frequency domain input signal.

- 30 **[0006]** The extension encoding unit may include a basic signal generator to generate the basic signal of the frequency domain input signal, using a frequency spectrum of the frequency domain input signal, a factor estimator to estimate an energy control factor using the basic signal, an energy extractor to extract an energy from the frequency domain input signal, an energy controller to control the extracted energy using the energy control factor, and an energy quantizer to quantize the controlled energy.

- 35 **[0007]** The basic signal generator may include an artificial signal generator to generate an artificial signal corresponding to a high-frequency section by copying and folding a low-frequency section of the frequency domain input signal, an envelope estimator to estimate an envelope of the artificial signal using a window, and an envelope applier to apply the estimated envelope to the artificial signal. Applying the estimated envelope means that the artificial signal is divided by the estimated envelope of the artificial signal.

- 40 **[0008]** The factor estimator may include a first tonality calculating unit to calculate a tonality of a high-frequency section of the frequency domain input signal, a second tonality calculating unit to calculate a tonality of the basic signal, and a factor calculating unit to calculate the energy control factor using the tonality of the high-frequency section and the tonality of the basic signal.

- [0009]** The foregoing and/or other aspects are also achieved by providing an encoding apparatus including a down-sampling unit to down-sample a time domain input signal, a core-encoding unit to core-encode the down-sampled time domain input signal, a frequency transforming unit to transform the core-encoded time domain input signal to a frequency domain input signal, and an extension encoding unit to perform bandwidth extension encoding using characteristics of the frequency domain input signal, and using a basic signal of the frequency domain input signal.

- 45 **[0010]** The extension encoding unit may include a basic signal generator to generate the basic signal of the frequency domain input signal, using a frequency spectrum of the frequency domain input signal, a factor estimator to estimate an energy control factor using the basic signal and the characteristics of the frequency domain input signal, an energy extractor to extract an energy from the frequency domain input signal, an energy controller to control the extracted energy using the energy control factor, and an energy quantizer to quantize the controlled energy.

- 50 **[0011]** The foregoing and/or other aspects are also achieved by providing an encoding apparatus including an encoding mode selecting unit to select an encoding mode of bandwidth extension encoding using a frequency domain input signal and a time domain input signal, and an extension encoding unit to perform the bandwidth extension encoding using the frequency domain input signal and the selected encoding mode.

**[0012]** The extension encoding unit may include an energy extractor to extract an energy from the frequency domain

input signal, based on the encoding mode, an energy controller to control the extracted energy based on the encoding mode, and an energy quantizer to quantize the controlled energy based on the encoding mode.

**[0013]** The foregoing and/or other aspects are achieved by providing a decoding apparatus including a core-decoding unit to core-decode a time domain input signal, the time domain input signal being contained in a bitstream and being core-encoded, an up-sampling unit to up-sample the core-decoded time domain input signal, a frequency transforming unit to transform the up-sampled time domain input signal to a frequency domain input signal, and an extension decoding unit to perform bandwidth extension decoding, using an energy of the time domain input signal and using the frequency domain input signal.

**[0014]** The extension decoding unit may include an inverse-quantizer to inverse-quantize the energy of the time domain input signal, a basic signal generator to generate a basic signal using the frequency domain input signal, a gain calculating unit to calculate a gain using the inverse-quantized energy and an energy of the basic signal, the gain being applied to the basic signal, and a gain applier to apply the calculated gain for each frequency band.

**[0015]** The basic signal generator may include an artificial signal generator to generate an artificial signal corresponding to a high-frequency section by copying and folding a low-frequency section of the frequency domain input signal, an envelope estimator to estimate an envelope of the basic signal using a window contained in the bitstream, and an envelope applier to apply the estimated envelope to the artificial signal.

**[0016]** The foregoing and/or other aspects are achieved by providing an encoding method including down-sampling a time domain input signal, core-encoding the down-sampled time domain input signal, transforming the time domain input signal to a frequency domain input signal, and performing bandwidth extension encoding using a basic signal of the frequency domain input signal.

**[0017]** The foregoing and/or other aspects are also achieved by providing an encoding method including down-sampling a time domain input signal, core-encoding the down-sampled time domain input signal, transforming the core-encoded time domain input signal to a frequency domain input signal, and performing bandwidth extension encoding using characteristics of the frequency domain input signal, and using a basic signal of the frequency domain input signal.

**[0018]** The foregoing and/or other aspects are also achieved by providing an encoding method including selecting an encoding mode of bandwidth extension encoding using a frequency domain input signal and a time domain input signal, and performing the bandwidth extension encoding using the frequency domain input signal and the selected encoding mode.

**[0019]** The foregoing and/or other aspects are achieved by providing a decoding method including core-decoding a time domain input signal, the time domain input signal being contained in a bitstream and being core-encoded, up-sampling the core-decoded time domain input signal, transforming the up-sampled time domain input signal to a frequency domain input signal, and performing bandwidth extension decoding, using an energy of the time domain input signal and using the frequency domain input signal.

**[0020]** Additional aspects, features, and/or advantages of example embodiments will be set forth in part in the description which follows and, in part, will be apparent from the description, or may be learned by practice of the disclosure.

**[0021]** According to example embodiments, a basic signal of an input signal may be extracted, and an energy of the input signal may be controlled using a tonality of a high-frequency domain of the input signal and using a tonality of the basic signal, and thus it is possible to efficiently extend a bandwidth of the high frequency domain.

#### Brief Description of Drawings

**[0022]** These and/or other aspects and advantages will become apparent and more readily appreciated from the following description of the example embodiments, taken in conjunction with the accompanying drawings of which:

FIG. 1 illustrates a block diagram of an encoding apparatus and a decoding apparatus according to example embodiments;

FIG. 2 illustrates a block diagram of an example of the encoding apparatus of FIG. 1;

FIG. 3 illustrates a block diagram of a core-encoding unit of the encoding apparatus of FIG. 1;

FIG. 4 illustrates a block diagram of an example of an extension encoding unit of the encoding apparatus of FIG. 1;

FIG. 5 illustrates a block diagram of another example of the extension encoding unit of the encoding apparatus of FIG. 1;

FIG. 6 illustrates a block diagram of a basic signal generator of the extension encoding unit;

FIG. 7 illustrates a block diagram of a factor estimator of the extension encoding unit;

FIG. 8 illustrates a flowchart of an operation of an energy quantizer of the encoding apparatus of FIG. 1;

FIG. 9 illustrates a diagram of an operation of quantizing an energy according to example embodiments;

FIG. 10 illustrates a diagram of an operation of generating an artificial signal according to example embodiments;

FIGS. 11A and 11B illustrate diagrams of examples of a window for estimating an envelope according to example embodiments;

FIG. 12 illustrates a block diagram of the decoding apparatus of FIG. 1;  
 FIG. 13 illustrates a block diagram of an extension decoding unit of FIG. 12;  
 FIG. 14 illustrates a flowchart of an operation of an inverse-quantizer of the extension decoding unit;  
 FIG. 15 illustrates a flowchart of an encoding method according to example embodiments;  
 FIG. 16 illustrates a flowchart of a decoding method according to example embodiments;  
 FIG. 17 illustrates a block diagram of another example of the encoding apparatus of FIG. 1;  
 FIG. 18 illustrates a block diagram of an operation of an energy quantizer of the encoding apparatus of FIG. 17;  
 FIG. 19 illustrates a diagram of an operation of quantizing an energy using an unequal bit allocation method according to example embodiments;  
 FIG. 20 illustrates a diagram of an operation of performing Vector Quantization (VQ) using intra frame prediction according to example embodiments;  
 FIG. 21 illustrates a diagram of an operation of quantizing an energy using a frequency weighting method according to example embodiments;  
 FIG. 22 illustrates a diagram of an operation of performing multi-stage split VQ, and VQ using intra frame prediction according to example embodiments;  
 FIG. 23 illustrates a block diagram of an operation of an inverse-quantizer of FIG. 13; and  
 FIG. 24 illustrates a block diagram of still another example of the encoding apparatus of FIG. 1.

#### Mode for the Invention

**[0023]** Reference will now be made in detail to example embodiments, examples of which are illustrated in the accompanying drawings, wherein like reference numerals refer to the like elements throughout. Example embodiments are described below to explain the present disclosure by referring to the figures.

**[0024]** FIG. 1 illustrates a block diagram of an encoding apparatus 101 and a decoding apparatus 102 according to example embodiments.

**[0025]** The encoding apparatus 101 may generate a basic signal of an input signal, and may transmit the generated basic signal to the decoding apparatus 102. Here, the basic signal may be generated based on a low-frequency signal, and may refer to a signal from which envelope information of the low-frequency signal is whitened and accordingly, the basic signal may be an excitation signal. When the basic signal is received, the decoding apparatus 102 may decode the input signal from the basic signal. In other words, the encoding apparatus 101 and the decoding apparatus 102 may perform Super Wide Band Bandwidth Extension (SWB BWE). Specifically, the SWB BWE may be performed to generate a signal in a high-frequency domain from 6.4 kilohertz (KHz) to 16 KHz corresponding to an SWB, based on a decoded Wide Band (WB) signal in a low-frequency domain from 0 KHz to 6.4 KHz. Here, the 16 KHz may vary depending on circumstances. Additionally, the decoded WB signal may be generated through a speech codec based on a Linear Prediction Domain (LPD)-based Code Excited Linear Prediction (CELP), or may be generated by a scheme of performing quantization in a frequency domain. The scheme of performing quantization in a frequency domain may include, for example, an Advanced Audio Coding (AAC) scheme performed based on Modified Discrete Cosine Transform (MDCT).

**[0026]** Hereinafter, operations of the encoding apparatus 101 and the decoding apparatus 102 will be further described.

**[0027]** FIG. 2 illustrates a block diagram of a configuration of the encoding apparatus 101 of FIG. 1.

**[0028]** Referring to FIG. 2, the encoding apparatus 101 may include, for example, a down-sampling unit 201, a core-encoding unit 202, a frequency transforming unit 203, and an extension encoding unit 204.

**[0029]** The down-sampling unit 201 may down-sample a time domain input signal for WB coding. Since the time domain input signal, namely an SWB signal, typically has a 32 KHz sampling rate, there is a need to convert the sampling rate into a sampling rate suitable for WB coding. For example, the down-sampling unit 201 may down-sample the time domain input signal from the 32 KHz sampling rate to a 12.8 KHz sampling rate.

**[0030]** The core-encoding unit 202 may core-encode the down-sampled time domain input signal. In other words, the core-encoding unit 202 may perform WB coding. For example, the core-encoding unit 202 may perform a CELP type WB coding.

**[0031]** The frequency transforming unit 203 may transform the time domain input signal to a frequency domain input signal. For example, the frequency transforming unit 203 may use either a Fast Fourier Transform (FFT) or an MDCT, to transform the time domain input signal to the frequency domain input signal. Hereinafter, it may be assumed that MDCT is applied.

**[0032]** The extension encoding unit 204 may perform bandwidth extension encoding using a basic signal of the frequency domain input signal. Specifically, the extension encoding unit 204 may perform SWB BWE encoding based on the frequency domain input signal.

**[0033]** Additionally, the extension encoding unit 204 may perform bandwidth extension encoding using characteristics of the frequency domain input signal and the basic signal of the frequency domain input signal. Here, the extension encoding unit 204 may be configured as illustrated in FIG. 4 or 5, depending on a source of the characteristics of the

frequency domain input signal.

**[0034]** An operation of the extension encoding unit 204 will be further described with reference to FIGS. 4 and 5 below.

**[0035]** In FIG. 2, an upper path indicates the core-encoding, and a lower path indicates the bandwidth extension encoding. In particular, energy information of the input signal may be transferred to the decoding apparatus 102 through the SWB BWE encoding.

**[0036]** FIG. 3 illustrates a block diagram of the core-encoding unit 202.

**[0037]** Referring to FIG. 3, the core-encoding unit 202 may include, for example, a signal classifier 301, and an encoder 302.

**[0038]** The signal classifier 301 may classify characteristics of the down-sampled input signal having the 12.8 KHz sampling rate. Specifically, the signal classifier 301 may determine an encoding mode to be applied to the frequency domain input signal, according to the characteristics of the frequency domain input signal. For example, in an International Telecommunications Union-Telecommunications (ITU-T) G.718 codec, the signal classifier 301 may determine a speech signal into one or more of a voiced speech encoding mode, a unvoiced speech encoding mode, a transient encoding mode, and a generic encoding mode. In this example, the unvoiced speech encoding mode may be designed to encode unvoiced speech frames and most of the inactive frames.

**[0039]** The encoder 302 may perform encoding optimized based on the characteristics of the frequency domain input signal classified by the signal classifier 301.

**[0040]** FIG. 4 illustrates a block diagram of an example of the extension encoding unit 204 of FIG. 2.

**[0041]** Referring to FIG. 4, the extension encoding unit 204 may include, for example, a basic signal generator 401, a factor estimator 402, an energy extractor 403, an energy controller 404, and an energy quantizer 405. In an example, the extension encoding unit 204 may estimate an energy control factor, without receiving an input of an encoding mode. In another example, the extension encoding unit 204 may estimate an energy control factor based on an encoding mode that is received from the core-encoding unit 202.

**[0042]** The basic signal generator 401 may generate a basic signal of an input signal using a frequency spectrum of the frequency domain input signal. The basic signal may refer to a signal used to perform SWB BWE based on a WB signal. In other words, the basic signal may refer to a signal used to form a fine structure of a low-frequency domain. An operation of generating a basic signal will be further described with reference to FIG. 6.

**[0043]** In an example, the factor estimator 402 may estimate an energy control factor using the basic signal. Specifically, the encoding apparatus 101 may transmit the energy information of the input signal to the decoding apparatus 102, in order to generate a signal in an SWB domain in the decoding apparatus 102. Additionally, the factor estimator 402 may estimate the energy control factor, to control the energy in a perceptual view. An operation of estimating the energy control factor will be further described with reference to FIG. 7.

**[0044]** In another example, the factor estimator 402 may estimate the energy control factor using the basic signal and the characteristics of the frequency domain input signal. In this example, the characteristics of the frequency domain input signal may be received from the core-encoding unit 202.

**[0045]** The energy extractor 403 may extract energy from the frequency domain input signal. The extracted energy may be transmitted to the decoding apparatus 102. Here, the energy may be extracted for each frequency band.

**[0046]** The energy controller 404 may control the extracted energy using the energy control factor. Specifically, the energy controller 404 may apply the energy control factor to the energy extracted for each frequency band, and may control the energy.

**[0047]** The energy quantizer 405 may quantize the controlled energy. The energy may be converted into a decibel (dB) scale, and may be quantized. Specifically, the energy quantizer 405 may acquire a global energy, namely a total energy, and may perform Scalar Quantization (SQ) on the global energy, and on a difference between the global energy and the energy for each frequency band. Additionally, a first band may directly quantize energy, and a following band may quantize a difference between a current band and a previous band. Furthermore, the energy quantizer 405 may directly quantize the energy for each frequency band, without using a difference value between frequency bands. When the energy is quantized for each frequency band, either SQ or Vector Quantization (VQ) may be used. The energy quantizer 405 will be further described with reference to FIGS. 8 and 9 below.

**[0048]** FIG. 5 illustrates a block diagram of another example of the extension encoding unit 204.

**[0049]** The extension encoding unit 204 of FIG. 5 may further include a signal classifier 501 and accordingly, may be different from the extension encoding unit 204 of FIG. 4. For example, the factor estimator 402 may estimate the energy control factor using the basic signal and the characteristics of the frequency domain input signal. In this example, the characteristics of the frequency domain input signal may be received from the signal classifier 501, instead of the core-encoding unit 202.

**[0050]** The signal classifier 501 may classify the input signal having the 32 KHz sampling rate based on the characteristics of the frequency domain input signal, using an MDCT spectrum. Specifically, the signal classifier 501 may determine an encoding mode to be applied to the frequency domain input signal, according to the characteristics of the frequency domain input signal.

**[0051]** When the characteristics of the input signal are classified, an energy control factor may be extracted from a signal and the energy may be controlled. In an embodiment, an energy control factor may only be extracted from a signal suitable for estimation of an energy control factor. For example, a signal that does not include a tonal component, such as a noise signal or unvoiced speech signal, may not be suitable for the estimation of the energy control factor. Here,

when the input signal is classified as the unvoiced speech encoding mode, the extension encoding unit 204 may perform bandwidth extension encoding, rather than estimating the energy control factor.

**[0052]** A basic signal generator 401, a factor estimator 402, an energy extractor 403, an energy controller 404, and an energy quantizer 405 shown in FIG. 5 may perform the same functions as the basic signal generator 401, the factor estimator 402, the energy extractor 403, the energy controller 404, and the energy quantizer 405 shown in FIG. 4, and accordingly further descriptions thereof will be omitted.

**[0053]** FIG. 6 illustrates a block diagram of the basic signal generator 401.

**[0054]** Referring to FIG. 6, the basic signal generator 401 may include, for example, an artificial signal generator 601, an envelope estimator 602, and an envelope applier 603.

**[0055]** The artificial signal generator 601 may generate an artificial signal corresponding to a high-frequency section by copying and folding a low-frequency section of the frequency domain input signal. Specifically, the artificial signal generator 601 may copy a low-frequency spectrum of the frequency domain input signal, and may generate an artificial signal in an SWB domain. An operation of generating an artificial signal will be further described with reference to FIG. 10.

**[0056]** The envelope estimator 602 may estimate an envelope of the basic signal using a window. The envelope of the basic signal may be used to remove envelope information of a low-frequency domain included in a frequency spectrum of the artificial signal in the SWB domain. An envelope of a predetermined frequency index may be determined using a frequency spectrum before and after the predetermined frequency. Additionally, an envelope may be estimated through a moving average. For example, when an MDCT is used to transform a frequency, the envelope of the basic signal may be estimated using an absolute value of an MDCT-transformed frequency spectrum.

**[0057]** Here, the envelope estimator 602 may form whitening bands, and may estimate an average of frequency magnitudes for each of the whitening bands as an envelope of a frequency contained in each of the whitening bands. A number of frequency spectrums contained in the whitening bands may be set to be less than a number of bands for extracting an energy.

**[0058]** When the average of frequency magnitudes for each of the whitening bands is estimated as the envelope of the frequency contained in each of the whitening bands, the envelope estimator 602 may transmit information including the number of frequency spectrums in the whitening bands, and may adjust a smoothness of the basic signal. Specifically, the envelope estimator 602 may transmit the information including the number of frequency spectrums in the whitening bands, based on whether a whitening band includes eight spectrums or three spectrums. For example, when a whitening band includes three spectrums, a further flattened basic signal may be generated, compared to a whitening band including eight spectrums.

**[0059]** Additionally, the envelope estimator 602 may estimate an envelope based on the encoding mode used during encoding by the core-encoding unit 202, rather than transmitting the information including the number of frequency spectrums in the whitening bands. The core-encoding unit 202 may classify the input signal into the voiced speech encoding mode, the unvoiced speech encoding mode, the transient encoding mode, and the generic encoding mode, based on the characteristics of the input signal, and may encode the input signal.

**[0060]** Here, the envelope estimator 602 may control the number of frequency spectrums contained in the whitening bands, based on the encoding modes according to the characteristics of the input signal. In an example, when the input signal is encoded based on the voiced speech encoding mode, the envelope estimator 602 may form a whitening band with three frequency spectrums, and may estimate an envelope. In another example, when the input signal is encoded based on encoding modes other than the voiced speech encoding mode, the envelope estimator 602 may form a whitening band with three frequency spectrums, and may estimate an envelope.

**[0061]** The envelope applier 603 may apply the estimated envelope to the artificial signal. An operation of applying the estimated envelope to the artificial signal is referred to as "whitening", and the artificial signal may be smoothed by the envelope. The envelope applier 603 may divide the artificial signal into envelopes of each frequency index, and may generate a basic signal.

**[0062]** FIG. 7 illustrates a block diagram of the factor estimator 402.

**[0063]** Referring to FIG. 7, the factor estimator 402 may include, for example, a first tonality calculating unit 701, a second tonality calculating unit 702, and a factor calculating unit 703.

**[0064]** The first tonality calculating unit 701 may calculate a tonality of a high-frequency section of the frequency domain input signal. In other words, the first tonality calculating unit 701 may calculate a tonality of an SWB domain, namely, the high-frequency section of the input signal.

**[0065]** The second tonality calculating unit 702 may calculate a tonality of the basic signal.

**[0066]** A tonality may be calculated by measuring a spectral flatness. Specifically, a tonality may be calculated using Equation 1 as below. The spectral flatness may be measured based on a relationship between a geometric average

and an arithmetic average of the frequency spectrum.

[Equation 1]

$$T = \min\left(10 \cdot \log_{10} \left( \frac{\prod_{k=0}^{N-1} |S(k)|^{\frac{1}{N}}}{\frac{1}{N} \sum_{k=0}^{N-1} |S(k)|} \right) / r, 0.999\right)$$

T: tonality, S(k): spectrum,

N: length of spectral coefficient,  $r$ : constant

**[0067]** The factor calculating unit 703 may calculate the energy control factor using the tonality of the high-frequency domain and the tonality of the basic signal. Here, the energy control factor may be calculated using the following Equation 2:

[Equation 2]

$$\alpha = \frac{N_o}{N_b} = \frac{(1 - T_o)}{(1 - T_b)}$$

$T_o$ : tonality of original spectrum,  $T_b$ : tonality of base spectrum

$N_o$ : noiseness factor of original spectrum,  $N_b$ : noiseness factor of base spectrum

**[0068]** In Equation 2,  $\alpha$  denotes an energy control factor,  $T_o$  denotes a tonality of an input signal, and  $T_b$  denotes a tonality of a basic signal. Additionally,  $N_b$  denotes a noiseness factor indicating how many noise components are contained in a signal.

The energy control factor may also be calculated using the following Equation 3:

[Equation 3]

$$\alpha = \frac{T_b}{T_o}$$

**[0069]** The factor calculating unit 703 may calculate the energy control factor for each frequency band. The calculated energy control factor may be applied to the energy of the input signal. Specifically, when the energy control factor is less than a predetermined energy control factor, the energy control factor may be applied to the energy of the input signal.

FIG. 8 illustrates a flowchart of an operation of the energy quantizer 405.

In operation 801, the energy quantizer 405 may pre-process an energy vector using the energy control factor, and may select a sub-vector of the pre-processed energy vector. For example, the energy quantizer 405 may subtract an average value from an energy value of each of selected energy vectors, or may calculate a weight for importance of each energy vector. Here, the weight for the importance may be calculated so that a quality of a complex sound may be maximized.

Additionally, the energy quantizer 405 may appropriately select a sub-vector of the energy vector, based on an encoding efficiency. To improve an interpolation effect, the energy quantizer 405 may select the sub-vector at regular intervals.

For example, the energy quantizer 405 may select a sub-vector based on the following Equation 4:

[Equation 4]

$$k^*n \ (n = 0, \dots, \text{and } N), k \geq 2$$

N is an integer less than a vector dimension.

**[0070]** In Equation 4, when k has a value of "2", only an even number may be selected as N.

In operation 802, the energy quantizer 405 may quantize and inverse-quantize the selected sub-vector. The energy quantizer 405 may select a quantization index for minimizing a Mean Square Error (MSE), and may quantize the selected sub-vector. Here, the MSE may be calculated using the following Equation 5:

[Equation 5]

$$MSE: d[x, y] = \frac{1}{N} \sum_{k=1}^N [x_k - y_k]^2$$

**[0071]** The energy quantizer 405 may quantize the sub-vector, based on one of SQ, VQ, Trellis Coded Quantization (TCQ), and Lattice Vector Quantization (LVQ). Here, the VQ may be performed based on either multi-stage VQ or split VQ, or may be performed using both the multi-stage VQ and split VQ. The quantization index may be transmitted to the decoding apparatus 102.

When the weight for the importance is calculated in operation 801, the energy quantizer 405 may obtain an optimized quantization index using a Weighted Mean Square Error (WMSE). Here, the WMSE may be calculated using the following Equation 6:

[Equation 6]

$$WMSE: d[x, y] = \frac{1}{N} \sum_{k=1}^N w_k [x_k - y_k]^2$$

In operation 803, the energy quantizer 405 may interpolate non-selected sub-vectors using the inverse-quantized sub-vector.

In operation 804, the energy quantizer 405 may calculate an interpolation error, namely, a difference between the interpolated non-selected sub-vectors and sub-vectors matched to the original energy vector.

In operation 805, the energy quantizer 405 may quantize the interpolation error. Here, the energy quantizer 405 may quantize the interpolation error using the quantization index for minimizing the MSE. The energy quantizer 405 may quantize the interpolation error based on one of the SQ, the VQ, the TCQ, and the LVQ. The VQ may be performed based on either multi-stage VQ or split VQ, or may be performed using both the multi-stage VQ and split VQ. When the weight for the importance is calculated in operation 801, the energy quantizer 405 may obtain an optimized quantization index using the WMSE.

**[0072]** In operation 806, the energy quantizer 405 may interpolate sub-vectors that are selected and quantized, may calculate the non-selected sub-vectors, and may add the interpolation error quantized in operation 805, to calculate a final quantized energy. Additionally, the energy quantizer 405 may perform post-processing to add the average value to the energy value, so that the final quantized energy may be obtained.

The energy quantizer 405 may perform multi-stage VQ using K candidates for the sub-vector, in order to improve a quantization performance using the same code book. For example, when at least two candidates for the sub-vector exist, the energy quantizer 405 may perform a distortion measure, and may determine an optimal candidate for the sub-vector. Here, the distortion measure may be determined based on two schemes.

In a first scheme, the energy quantizer 405 may generate an index set for minimizing an MSE or WMSE in each stage for each candidate, and may select candidates for a sub-vector having a smallest sum of an MSE or WMSE in all stages. Here, the first scheme may have an advantage of a simple calculation.

In a second scheme, the energy quantizer 405 may generate an index set for minimizing an MSE or WMSE in each stage for each candidate, may restore the energy vector through an inverse-quantization operation, and may select candidates for a sub-vector for minimizing an MSE or WMSE between the restored energy vector and an original energy vector. Here, the MSE may be obtained using an actual quantized value, even when a calculation amount for restoration is added. Thus, the second scheme may have an advantage of an excellent performance.

FIG. 9 illustrates an operation of quantizing an energy according to example embodiments.

Referring to FIG. 9, an energy vector may represent 14 dimensions. In a first stage of FIG. 9, the energy quantizer 405 may select only even numbers from the energy vector, and may select a sub-vector corresponding to 7 dimensions. In a second stage, the energy quantizer 405 may perform VQ that is split into two quantization stages.

In the second stage, the energy quantizer 405 may perform quantization using an error signal of the first stage. The energy quantizer 405 may obtain an interpolation error through an operation of inverse-quantizing the selected sub-



vector. In a third stage, the energy quantizer 405 may quantize the interpolation error through two split VQ.

FIG. 10 illustrates a diagram of an operation of generating an artificial signal according to example embodiments.

Referring to FIG. 10, the artificial signal generator 601 may copy a frequency spectrum 1001 corresponding to a low-frequency domain from fL KHz to 6.4 KHz in a total frequency band. The copied frequency spectrum 1001 may be shifted to a frequency domain from 6.4 KHz to 12.8-fL KHz. Additionally, a frequency spectrum corresponding to a frequency domain from 12.8-fL KHz to 16 KHz may be generated by folding a frequency spectrum corresponding to the frequency domain from 6.4 KHz to 12.8-fL KHz. In other words, an artificial signal corresponding to an SWB domain, namely a high-frequency domain, may be generated in a frequency domain from 6.4 KHz to 16 KHz.

Here, when an MDCT is used to generate a frequency spectrum, a relationship between fL KHz and 6.4 KHz may exist. Specifically, when a frequency index of the MDCT corresponding to 6.4 KHz is an even number, a frequency index for fL KHz may need to be an even number. Conversely, when the frequency index of the MDCT corresponding to 6.4 KHz is an odd number, the frequency index for fL KHz may need to be an odd number.

For example, when an MDCT is applied to extract 640 spectrums for the original input signal, a 256-th frequency index may correspond to 6.4 KHz, and the frequency index of the MDCT corresponding to 6.4 KHz may be an even number (6400/16000\*640). In this example, fL needs to be selected as an even number. In other words, 2 (50 Hz), 4 (100 Hz) and the like may be used as fL. The operation of FIG. 10 may be equally applied to a decoding operation.

FIGS. 11A and 11B illustrate diagrams of examples of a window for estimating an envelope according to example embodiments.

Referring to FIGS. 11A and 11B, a peak of a window 1101 and a peak of a window 1102 may each indicate a frequency index where a current envelope is to be estimated. The envelope of the basic signal may be estimated using the following Equation 7:

[Equation 7]

$$Env(n) = \sum_{k=n-d}^{n+d} w(k - n + d) \cdot |S(k)|$$

*Env(n)*: Envelope, *w(k)*: window, *S(k)*: Spectrum, *n*: frequency index,  
2d+1: window length

**[0073]** The windows 1101 and 1102 may be used to be fixed at all times, and there is no need to additionally transmit a bit. When the windows 1101 and 1102 are selectively used, information indicating which window is used to estimate an envelope may be represented by bits, and may be additionally transferred to the decoding apparatus 102. The bits may be transmitted for each frequency band, or may be transmitted to a single frame all at once.

Comparing the windows 1101 and 1102, the window 1102 may be used to estimate an envelope by further applying a weight to a frequency spectrum corresponding to a current frequency index, compared with the window 1101. Accordingly, a basic signal generated by the window 1102 may be smoother than a basic signal generated by the window 1101. A type of window may be selected by comparing a frequency spectrum of an input signal with a frequency spectrum of a basic signal generated by the window 1101 or window 1102. Additionally, a window enabling similar tonality through comparison of a tonality of a high-frequency section may be selected. Moreover, a window having a high correlation may be selected by comparing a correlation of high-frequency sections.

FIG. 12 illustrates a block diagram of the decoding apparatus 102 of FIG. 1.

The decoding apparatus 102 of FIG. 12 may perform an operation inverse to the encoding apparatus 101 of FIG. 2.

Referring to FIG. 12, the decoding apparatus 102 may include, for example, a core-decoding unit 1201, an up-sampling unit 1202, a frequency transforming unit 1203, an extension decoding unit 1204, and an inverse frequency transforming unit 1205.

The core-decoding unit 1201 may core-decode a time domain input signal that is included in a bitstream and that is core-encoded. A signal with a 12.8 KHz sampling rate may be extracted through the core-decoding.

The up-sampling unit 1202 may up-sample the core-decoded time domain input signal. A signal with a 32 KHz sampling rate may be extracted through the up-sampling.

The frequency transforming unit 1203 may transform the up-sampled time domain input signal to a frequency domain input signal. The up-sampled time domain input signal may be transformed using the same scheme as the frequency transformation scheme used by the encoding apparatus 101, for example, an MDCT scheme may be used.

The extension decoding unit 1204 may perform bandwidth extension decoding using an energy of the time domain input signal and using the frequency domain input signal. An operation of the extension decoding unit 1204 will be further described with reference to FIG. 13.

The inverse frequency transforming unit 1205 may perform inverse frequency transformation with respect to a result of

the bandwidth extension decoding. Here, the inverse frequency transformation may be performed in a manner inverse to the frequency transformation scheme used by the frequency transforming unit 1203. For example, the inverse frequency transforming unit 1205 may perform an Inverse Modified Discrete Cosine Transform (IMDCT).

FIG. 13 illustrates a block diagram of the extension decoding unit 1204 of FIG. 12.

Referring to FIG. 13, the extension decoding unit 1204 may include, for example, an inverse-quantizer 1301, a gain calculating unit 1302, a gain applier 1303, an artificial signal generator 1304, an envelope estimator 1305, and an envelope applier 1306.

The inverse-quantizer 1301 may inverse-quantize the energy of the time domain input signal. An operation of inverse-quantizing the energy will be further described with reference to FIG. 14.

The gain calculating unit 1302 may calculate a gain to be applied to the basic signal, using the inverse-quantized energy and an energy of the basic signal. Specifically, the gain may be determined based on a ratio of the inverse-quantized energy and the energy of the basic signal. Since an energy is typically determined based on a sum of squares of an amplitude of each frequency spectrum, a root value of an energy ratio may be used.

The gain applier 1303 may apply the calculated gain for each frequency band. Accordingly, a frequency spectrum of an SWB may be finally determined.

In an example, the calculating and applying of the gain may be performed by matching a band to a band used to transmit energy, as described above. In another example, to prevent a rapid change in energy, the gain may be calculated and applied by dividing an overall frequency band into sub-bands. In this example, an inverse-quantized energy of a neighboring band may be interpolated, and an energy in a band boundary may be smoothed. For example, each band may be divided into three sub-bands, and an inverse-quantized energy of a current band may be allocated to an intermediate sub-band among the three sub-bands. Subsequently, gains of a first sub-band and a third sub-band may be calculated using a newly smoothed energy, based on an energy allocated to an intermediate band between a previous band and a next band, and based on interpolation. In other words, the gain may be calculated for each band.

Such an energy smoothing scheme may be applied to be fixed at all times. Additionally, the extension encoding unit 204 may transmit information indicating that the energy smoothing scheme is required, and may apply the energy smoothing scheme to only frames requiring the energy smoothing scheme. Here, when smoothing is performed and when less quantization error of a total energy occurs, information indicating a frame requiring the energy smoothing scheme may be selected, compared to when the smoothing is not performed.

A basic signal may be generated using the frequency domain input signal. An operation of generating a basic signal may be performed using components as described below.

The artificial signal generator 1304 may generate an artificial signal corresponding to a high-frequency section by copying and folding a low-frequency section of the frequency domain input signal. Here, the frequency domain input signal may be a WB-decoded signal with a 32 KHz sampling rate.

The envelope estimator 1305 may estimate an envelope of the basic signal using a window contained in the bitstream.

The window may be used to estimate the envelope by the encoding apparatus 101. A type of window may be bit type, and the window may be contained in a bitstream and may be transmitted to the decoding apparatus 102.

The envelope applier 1306 may apply the estimated envelope to the artificial signal, and may generate a basic signal. For example, when the average of frequency magnitudes for each of the whitening bands is estimated as the envelope of the frequency contained in each of the whitening bands, the envelope estimator 602 of the encoding apparatus 101 may transmit, to the decoding apparatus 102, the information including the number of frequency spectrums in the whitening bands. When the information is received, the envelope estimator 1305 of the decoding apparatus 102 may estimate an envelope based on the received information, and the envelope applier 1306 may apply the estimated envelope. Additionally, the envelope estimator 1305 may estimate an envelope based on a core-decoding mode used by the core-decoding unit 1201, rather than transmitting the information including the number of frequency spectrums in the whitening bands.

The core-decoding unit 1201 may determine a decoding mode among a voiced speech decoding mode, an unvoiced speech decoding mode, a transient decoding mode, and a generic decoding mode, based on characteristics of a frequency domain input signal, and may perform decoding in the determined decoding mode. Here, the envelope estimator 1305 may control the number of frequency spectrums in the whitening bands, using the decoding mode based on the characteristics of the frequency domain input signal. In an example, when the frequency domain input signal is decoded in the voiced speech decoding mode, the envelope estimator 1305 may form a whitening band with three frequency spectrums, and may estimate an envelope. In another example, when the frequency domain input signal is decoded in decoding modes other than the voiced speech decoding mode, the envelope estimator 1305 may form a whitening band with three frequency spectrums, and may estimate an envelope.

FIG. 14 illustrates a flowchart of an operation of the inverse-quantizer 1301.

In operation 1401, the inverse-quantizer 1301 may inverse-quantize the selected sub-vector of the energy vector, using an index 1 received from the encoding apparatus 101.

**[0074]** In operation 1402, the inverse-quantizer 1301 may inverse-quantize an interpolation error corresponding to

non-selected sub-vectors, using an index 2 received from the encoding apparatus 101.

In operation 1403, the inverse-quantizer 1301 may interpolate the inverse-quantized sub-vector, and may calculate the non-selected sub-vectors. Additionally, the inverse-quantizer 1301 may add the inverse-quantized interpolation error to the non-selected sub-vectors. Furthermore, the inverse-quantizer 1301 may perform post-processing to add an average value that is subtracted in a preprocessing operation, and may calculate a final inverse-quantized energy.

FIG. 15 illustrates a flowchart of an encoding method according to example embodiments.

In operation 1501, the encoding apparatus 101 may down-sample a time domain input signal.

In operation 1502, the encoding apparatus 101 may core-encode the down-sampled time domain input signal.

In operation 1503, the encoding apparatus 101 may transform the time domain input signal to a frequency domain input signal.

In operation 1504, the encoding apparatus 101 may perform bandwidth extension encoding on the frequency domain input signal. For example, the encoding apparatus 101 may perform the bandwidth extension encoding based on encoding information determined in operation 1502. Here, the encoding information may include an encoding mode classified based on characteristics of the frequency domain input signal.

For example, the encoding apparatus 101 may perform the bandwidth extension encoding by the following operations. The encoding apparatus 101 may generate a basic signal of the frequency domain input signal, using a frequency spectrum of the frequency domain input signal. Also, the encoding apparatus 101 may generate a basic signal of the frequency domain input signal, using characteristics of the frequency domain input signal and a frequency spectrum of the frequency domain input signal. Here, the characteristics of the frequency domain input signal may be derived through core-encoding, or a separate signal classification. Additionally, the encoding apparatus 101 may estimate an energy control factor using the basic signal. Subsequently, the encoding apparatus 101 may extract an energy from the frequency domain input signal. The encoding apparatus 101 may control the extracted energy using the energy control factor. The encoding apparatus 101 may quantize the controlled energy.

Here, the basic signal may be generated through the following schemes:

The encoding apparatus 101 may generate an artificial signal corresponding to a high-frequency section by copying and folding a low-frequency section of the frequency domain input signal. Additionally, the encoding apparatus 101 may estimate an envelope of the basic signal using a window. Here, the encoding apparatus 101 may select a window based on a comparison result of either a tonality or a correlation, and may estimate the envelope of the basic signal. For example, the encoding apparatus 101 may estimate an average of frequency magnitudes in each of whitening bands, as an envelope of a frequency contained in each of the whitening bands. Specifically, the encoding apparatus 101 may control a number of frequency spectrums in each of the whitening bands, based on a core-encoding mode, and may estimate the envelope of the basic signal.

Subsequently, the encoding apparatus 101 may apply the estimated envelope to the artificial signal, so that the basic signal may be generated.

The energy control factor may be estimated using the following scheme:

The encoding apparatus 101 may calculate a tonality of a high-frequency section of the frequency domain input signal. Additionally, the encoding apparatus 101 may calculate a tonality of the basic signal. Subsequently, the encoding apparatus 101 may calculate the energy control factor using the tonality of the high-frequency section and the tonality of the basic signal.

Additionally, the energy may be quantized through the following scheme:

The encoding apparatus 101 may select a sub-vector of an energy vector, may quantize the selected sub-vector, and may quantize non-selected sub-vectors using an interpolation error. Here, the encoding apparatus 101 may select a sub-vector at regular intervals.

For example, the encoding apparatus 100 may select candidates for the sub-vector, and may perform multi-stage VQ including at least two stages. In this example, the encoding apparatus 100 may generate an index set for minimizing an MSE or WMSE in each stage for each of the candidates for the sub-vector, and may select candidates for a sub-vector having a smallest sum of an MSE or WMSE in all stages. Alternatively, the encoding apparatus 100 may generate an index set for minimizing an MSE or a WMSE in each stage for each of the candidates for the sub-vector, may restore the energy vector through an inverse-quantization operation, and may select candidates for a sub-vector for minimizing an MSE or WMSE between the restored energy vector and an original energy vector.

FIG. 16 illustrates a flowchart of a decoding method according to example embodiments.

In operation 1601, the decoding apparatus 102 may core-decode a time domain input signal that is included in a bitstream and that is core-encoded.

In operation 1602, the decoding apparatus 102 may up-sample the core-decoded time domain input signal.

In operation 1603, the decoding apparatus 102 may transform the up-sampled time domain input signal to a frequency domain input signal.

In operation 1604, the decoding apparatus 102 may perform bandwidth extension decoding using an energy of the time domain input signal and using the frequency domain input signal.

Specifically, the bandwidth extension decoding may be performed as below.

The decoding apparatus 102 may inverse-quantize the energy of the time domain input signal. Here, the decoding apparatus 102 may select a sub-vector of an energy vector, may inverse-quantize the selected sub-vector, may interpolate the inverse-quantized sub-vector, and may add an interpolation error to the interpolated sub-vector, to finally inverse-quantize the energy.

Additionally, the decoding apparatus 102 may generate a basic signal using the frequency domain input signal. Subsequently, the decoding apparatus 102 may calculate a gain to be applied to the basic signal, using the inverse-quantized energy and an energy of the basic signal. Finally, the decoding apparatus 102 may apply the calculated gain for each frequency band.

Specifically, the basic signal may be generated as below.

**[0075]** The decoding apparatus 102 may generate an artificial signal corresponding to a high-frequency section by copying and folding a low-frequency section of the frequency domain input signal. Additionally, the decoding apparatus 102 may estimate an envelope of the basic signal using a window contained in the bitstream. Here, when window information is set to be equally used, the window may not be contained in the bitstream. Subsequently, the decoding apparatus 102 may apply the estimated envelope to the artificial signal.

Other descriptions of FIGS. 15 and 16 have been already given above with reference to FIGS. 1 through 14.

FIG. 17 illustrates a block diagram of another example of the encoding apparatus 100 according to example embodiments. Referring to FIG. 17, the encoding apparatus 100 may include, for example, an encoding mode selecting unit 1701, and an extension encoding unit 1702.

The encoding mode selecting unit 1701 may select an encoding mode of bandwidth extension encoding using a frequency domain input signal and a time domain input signal.

Specifically, the encoding mode selecting unit 1701 may classify a frequency domain input signal using the frequency domain input signal and the time domain input signal, may determine the encoding mode of the bandwidth extension encoding mode, and may determine a number of frequency bands based on the determined encoding mode. Here, to improve a performance of the extension encoding unit 1702, the encoding mode may be set as a set of an encoding mode determined during core-encoding, and another encoding mode.

The encoding mode may be classified, for example, into a normal mode, a harmonic mode, a transient mode, and a noise mode. First, the encoding mode selecting unit 1701 may determine whether a current frame is a transient frame, based on a ratio of a long-term energy of the time domain input signal to a high-band energy of the current frame. A transient signal interval may refer to an interval where energy is rapidly changed in a time domain, that is, an interval where the high-band energy is rapidly changed.

The normal mode, the harmonic mode, and the noise mode may be determined as follows: First, the encoding mode selecting unit 1701 may acquire a global energy of a frequency domain of a previous frame and a current frame, may divide a ratio of the global energies and the frequency domain input signal by a frequency band defined in advance, and may determine the normal mode, the harmonic mode, and the noise mode using an average energy and a peak energy of each frequency band. The harmonic mode may provide a signal having a largest difference between an average energy and a peak energy in a frequency domain signal. The noise mode may provide a signal having a small change in energy. The normal mode may provide signals other than the signal of the harmonic mode and the signal of the noise mode.

Additionally, a number of frequency bands in the normal mode and the harmonic mode may be determined to be "16", and a number of frequency bands in the transient mode may be determined to be "5". Furthermore, a number of frequency bands in the noise mode may be determined to be "12". The extension encoding unit 1702 may perform the bandwidth extension encoding using the frequency domain input signal and the encoding mode. Referring to FIG. 17, the extension encoding unit 1702 may include, for example, a basic signal generator 1703, a factor estimator 1704, an energy extractor 1705, an energy controller 1706, and an energy quantizer 1707. The basic signal generator 1703 and the factor estimator 1704 may perform the same functions as the basic signal generator 401 and the factor estimator 402 of FIG. 4 and accordingly, further descriptions thereof will be omitted.

The energy extractor 1705 may extract an energy corresponding to each frequency band, based on the number of frequency bands determined depending on the encoding mode. The energy controller 1706 may control the extracted energy based on the encoding mode.

The basic signal generator 1703, the factor estimator 1704, and the energy controller 1706 may be used or not be used, based on the encoding mode. For example, in the normal mode and the harmonic mode, the basic signal generator 1703, the factor estimator 1704, and the energy controller 1706 may be used, however, in the transient mode and the noise mode, the basic signal generator 1703, the factor estimator 1704, and the energy controller 1706 may not be used. Further descriptions of the basic signal generator 1703, the factor estimator 1704, and the energy controller 1706 have been given above with reference to FIG. 4.

The energy quantizer 1707 may quantize the energy controlled based on the encoding mode. In other words, a band energy passing through an energy control operation may be quantized by the energy quantizer 1707.

FIG. 18 illustrates a diagram of an operation performed by the energy quantizer 1707.

The energy quantizer 1707 may quantize an energy extracted from the frequency domain input signal, based on the encoding mode. Here, the energy quantizer 1707 may quantize a band energy using a scheme optimized for each input signal, based on perceptual characteristics of each input signal and the number of frequency bands, depending on the encoding mode.

In an example, when the transient mode is used as the encoding mode, the energy quantizer 1707 may quantize five band energies using a frequency weighting method based on the perceptual characteristics. In another example, when the normal mode or the harmonic mode is used as the encoding mode, the energy quantizer 1707 may quantize 16 band energies using an unequal bit allocation method based on the perceptual characteristics. When the perceptual characteristics are unclear, the energy quantizer 1707 may perform typical quantization, regardless of the perceptual characteristics.

FIG. 19 illustrates a diagram of an operation of quantizing an energy using the unequal bit allocation method according to example embodiments.

The unequal bit allocation method may be performed based on perceptual characteristics of an input signal targeted for extension encoding, and be used to more accurately quantize a band energy corresponding to a lower frequency band having a high perceptual importance. Accordingly, the energy quantizer 1707 may allocate, to the band energy corresponding to the lower frequency band, a number of bits that are equal to or greater than a number of band energies, and may determine the perceptual importance of the band energy.

For example, the energy quantizer 1707 may allocate a greater number of bits to lower frequency bands 0 to 5, so that a same number of bits may be allocated to the lower frequency bands 0 to 5. Additionally, as a frequency band increases, a number of bits allocated by the energy quantizer 1707 to the frequency band decreases. Accordingly, a bit allocation may enable frequency bands 0 to 13 to be quantized as shown in FIG. 19, and may enable frequency bands 14 and 15 to be quantized as shown in FIG. 20.

FIG. 20 illustrates a diagram of an operation of performing VQ using intra frame prediction according to example embodiments.

The energy quantizer 1707 may predict a representative value of a quantization target vector having at least two elements, and may perform VQ on an error signal between the predicted representative value and at least two elements of the quantization target vector.

Such an intra frame prediction may be shown in FIG. 20, and a scheme of predicting a representative value of a quantization target vector and deriving an error signal may be represented by the following Equation 8:

[Equation 8]

$$\begin{aligned} p &= 0.4 * QEnv(12) + 0.6 * QEnv(13) \\ e(14) &= Env(14) - p \\ e(15) &= Env(15) - p \end{aligned}$$

In Equation 8, Env(n) denotes a non-quantized band energy, and QEnv(n) denotes a quantized band energy. Additionally, p denotes the predicted representative value of the quantization target vector, and e(n) denotes an error energy. Here, VQ may be performed on e(14) and e(15).

FIG. 21 illustrates a diagram of an operation of quantizing an energy using the frequency weighting method according to example embodiments.

The frequency weighting method may be used to more accurately quantize a band energy corresponding to a lower frequency band having a high perceptual importance, based on perceptual characteristics of an input signal targeted for extension encoding, in the same manner as the unequal bit allocation method. Accordingly, the energy quantizer 1707 may allocate, to the band energy corresponding to the lower frequency band, a number of bits that are equal to or greater than a number of band energies, and may determine the perceptual importance.

For example, the energy quantizer 1707 may assign a weight of "1.0" to a band energy corresponding to frequency bands 0 to 3, namely lower frequency bands, and may assign a weight of "0.7" to a band energy corresponding to a frequency band 15, namely a higher frequency band. To use the assigned weights, the energy quantizer 1707 may obtain an optimal index using a WMSE value.

FIG. 22 illustrates a diagram of an operation of performing multi-stage split VQ, and VQ using intra frame prediction according to example embodiments.

The energy quantizer 1707 may perform VQ on the normal mode with 16 band energies, as shown in FIG. 22. Here, the energy quantizer 1707 may perform the VQ using the unequal bit allocation method, the intra frame prediction, and the multi-stage split VQ with energy interpolation.

FIG. 23 illustrates a diagram of an operation performed by the inverse-quantizer 1301.

The operation of FIG. 23 may be performed in an inverse manner to the operation of FIG. 18. When an encoding mode is used during extension encoding, as shown in FIG. 17, the inverse-quantizer 1301 of the extension decoding unit 1204 may decode the encoding mode.

The inverse-quantizer 1301 may decode the encoding mode using an index that is received first. Subsequently, the inverse-quantizer 1301 may perform inverse-quantization using a scheme set based on the decoded encoding mode. Referring to FIG. 23, the inverse-quantizer 1301 may inverse-quantize blocks respectively corresponding to encoding modes, in an inverse order of the quantization. An energy vector quantized using the Multi-stage split VQ with energy interpolation may be inverse-quantized in the same manner as shown in FIG. 14. In other words, the inverse-quantizer 1301 may perform inverse-quantization using the intra frame prediction, through the following Equation 9:

[Equation 9]

$$\begin{aligned} p &= 0.4 * QEnv(12) + 0.6 * QEnv(13) \\ QEnv(14) &= \hat{e}(14) + p \\ QEnv(15) &= \hat{e}(15) + p \end{aligned}$$

In Equation 9, Env(n) denotes a non-quantized band energy, and QEnv(n) denotes a quantized band energy. Additionally, p denotes the predicted representative value of the quantization target vector, and  $\hat{e}(n)$  denotes a quantized error energy.

**[0076]** FIG. 24 illustrates a block diagram of still another example of the encoding apparatus 101.

The encoding apparatus 101 of FIG. 24 may include, for example, a down-sampling unit 2401, a core-encoding unit 2402, a frequency transforming unit 2403, and an extension encoding unit 2404.

The down-sampling unit 2401, the core-encoding unit 2402, the frequency transforming unit 2403, and the extension encoding unit 2404 in the encoding apparatus 101 of FIG. 24 may perform the same basic operations as the down-sampling unit 201, the core-encoding unit 202, the frequency transforming unit 203, and the extension encoding unit 204 in the encoding apparatus 101 of FIG. 2. However, the extension encoding unit 2404 need not transmit information to the core-encoding unit 2402, and may directly receive a time domain input signal.

The methods according to the above-described example embodiments may be recorded in non-transitory computer-readable media including program instructions to implement various operations embodied by a computer. The media may also include, alone or in combination with the program instructions, data files, data structures, and the like. The program instructions recorded on the media may be those specially designed and constructed for the purposes of the example embodiments, or they may be of the kind well-known and available to those having skill in the computer software arts. Examples of non-transitory computer-readable media include magnetic media such as hard disks, floppy disks, and magnetic tape; optical media such as CD ROM disks and DVDs; magneto-optical media such as optical disks; and hardware devices that are specially configured to store and perform program instructions, such as read-only memory (ROM), random access memory (RAM), flash memory, and the like.

Examples of program instructions include both machine code, such as produced by a compiler, and files containing higher level code that may be executed by the computer using an interpreter. The described hardware devices may be configured to act as one or more software modules in order to perform the operations of the above-described example embodiments, or vice versa. Any one or more of the software modules described herein may be executed by a dedicated processor unique to that unit or by a processor common to one or more of the modules. The described methods may be executed on a general purpose computer or processor or may be executed on a particular machine such as the encoding apparatuses and decoding apparatuses described herein.

Although example embodiments have been shown and described, it would be appreciated by those skilled in the art that changes may be made in these example embodiments without departing from the principles and spirit of the disclosure, the scope of which is defined in the claims and their equivalents.

**[0077]** Attention is directed to all papers and documents which are filed concurrently with or previous to this specification in connection with this application and which are open to public inspection with this specification, and the contents of all such papers and documents are incorporated herein by reference.

**[0078]** All of the features disclosed in this specification (including any accompanying claims, abstract and drawings), and/or all of the steps of any method or process so disclosed, may be combined in any combination, except combinations where at least some of such features and/or steps are mutually exclusive.

**[0079]** Each feature disclosed in this specification (including any accompanying claims, abstract and drawings) may be replaced by alternative features serving the same, equivalent or similar purpose, unless expressly stated otherwise. Thus, unless expressly stated otherwise, each feature disclosed is one example only of a generic series of equivalent or similar features.

**[0080]** The invention is not restricted to the details of the foregoing embodiment(s). The invention extends to any novel one, or any novel combination, of the features disclosed in this specification (including any accompanying claims, abstract and drawings), or to any novel one, or any novel combination, of the steps of any method or process so disclosed.

**[0081]** The invention might include, relate to, and/or be defined by, the following aspects:

[Aspect 1] An encoding apparatus, comprising: a processor to control one or more processor-executable units; a core-encoding unit to core-encode a down-sampled time domain input signal; a frequency transforming unit to transform the time domain input signal to a frequency domain input signal; and an extension encoding unit to perform bandwidth extension encoding using a basic signal of the frequency domain input signal and the frequency domain input signal.

[Aspect 2] The encoding apparatus of Aspect 1, wherein the extension encoding unit comprises: a basic signal generator to generate the basic signal of the frequency domain input signal using a frequency spectrum of the frequency domain input signal; a factor estimator to estimate an energy control factor using the basic signal and the frequency domain input signal; an energy extractor to extract an energy from the frequency domain input signal; an energy controller to control the extracted energy using the energy control factor; and an energy quantizer to quantize the controlled energy.

[Aspect 3] The encoding apparatus of Aspect 2, wherein the basic signal generator comprises: an artificial signal generator to generate an artificial signal corresponding to a high-frequency section by copying and folding a low-frequency section of the frequency domain input signal; an envelope estimator to estimate an envelope of the artificial signal using a window; and an envelope applier to apply the estimated envelope to the artificial signal for generating the basic signal.

[Aspect 4] The encoding apparatus of Aspect 3, wherein a peak of the window corresponds to a frequency index where an envelope of the basic signal is to be estimated, and wherein the envelope estimator selects a window based on a comparison result of either a tonality or a correlation, and estimates the envelope of the basic signal using the selected window.

[Aspect 5] The encoding apparatus of Aspect 3, wherein the envelope estimator estimates an average of frequency magnitudes of a whitening band as an envelope of a frequency contained in the whitening band.

[Aspect 6] The encoding apparatus of Aspect 5, wherein the envelope estimator controls a number of frequency spectrums contained in the whitening band based on a core-encoding mode, and estimates the envelope of the basic signal.

[Aspect 7] The encoding apparatus of Aspect 2, wherein the factor estimator comprises: a first tonality calculating unit to calculate a tonality of a high-frequency section of the frequency domain input signal; a second tonality calculating unit to calculate a tonality of the basic signal; and a factor calculating unit to calculate the energy control factor using the tonality of the high-frequency section and the tonality of the basic signal.

[Aspect 8] The encoding apparatus of Aspect 2, wherein the energy controller controls the extracted energy when the energy control factor is less than a predetermined energy control factor.

[Aspect 9] The encoding apparatus of Aspect 2, wherein the energy quantizer selects a sub-vector of an energy vector, quantizes the selected sub-vector, and quantizes non-selected sub-vectors using an interpolation error.

[Aspect 10] The encoding apparatus of Aspect 9, wherein the energy quantizer selects the sub-vector and quantizes the selected sub-vector at regular intervals.

[Aspect 11] The encoding apparatus of Aspect 9, wherein the energy quantizer selects candidates for the sub-vector, and performs multi-stage Vector Quantization (VQ) including at least two stages.

[Aspect 12] The encoding apparatus of Aspect 9, wherein the energy quantizer generates an index set for minimizing a Mean Square Error (MSE) or a Weighted Mean Square Error (WMSE) in each stage for each of the candidates for the sub-vector, and selects candidates for a sub-vector having a smallest sum of an MSE or WMSE in all stages.

[Aspect 13] The encoding apparatus of Aspect 9, wherein the energy quantizer generates an index set for minimizing an MSE or a WMSE in each stage for each of the candidates for the sub-vector, restores an energy vector through an inverse-quantization operation, and selects candidates for a sub-vector for minimizing an MSE or WMSE between the restored energy vector and an original energy vector.

[Aspect 14] An encoding apparatus, comprising: a processor to control one or more processor-executable units; a core-encoding unit to core-encode a down-sampled time domain input signal; a frequency transforming unit to transform the time domain input signal to a frequency domain input signal; and an extension encoding unit to perform bandwidth extension encoding using the frequency domain input signal, and using a basic signal of the frequency domain input signal.

[Aspect 15] The encoding apparatus of Aspect 14, wherein the extension encoding unit comprises: a basic signal generator to generate the basic signal of the frequency domain input signal using a frequency spectrum of the frequency domain input signal; a factor estimator to estimate an energy control factor using the characteristics of the basic signal and the frequency domain input signal; an energy extractor to extract an energy from the frequency

domain input signal; an energy controller to control the extracted energy using the energy control factor; and an energy quantizer to quantize the controlled energy.

[Aspect 16] The encoding apparatus of Aspect 15, wherein the extension encoding unit further comprises a signal classifier to classify the frequency domain input signal based on characteristics of the frequency domain input signal, using the frequency spectrum of the frequency domain input signal, and wherein the factor estimator estimates the energy control factor using the characteristics of the frequency domain input signal determined by the signal classifier.

[Aspect 17] The encoding apparatus of Aspect 15, wherein the factor estimator estimates the energy control factor using the characteristics of the frequency domain input signal determined by the core-encoding unit.

[Aspect 18] The encoding apparatus of Aspect 15, wherein the basic signal generator comprises: an artificial signal generator to generate an artificial signal corresponding to a high-frequency section by copying and folding a low-frequency section of the input signal of frequency domain; an envelope estimator to estimate an envelope of the artificial signal using a window; and an envelope applier to apply the estimated envelope to the artificial signal for generating the basic signal.

[Aspect 19] The encoding apparatus of Aspect 18, wherein a peak of the window corresponds to a frequency index where an envelope of the basic signal is to be estimated, and wherein the envelope estimator selects a window based on a comparison result of either a tonality or a correlation, and estimates the envelope of the basic signal using selected the window.

[Aspect 20] The encoding apparatus of Aspect 18, wherein the envelope estimator estimates an average of frequency magnitudes of a whitening band as an envelope of a frequency contained in the whitening band.

[Aspect 21] The encoding apparatus of Aspect 20, wherein the envelope estimator controls a number of frequency spectrums contained in the whitening band based on a core-encoding mode, and estimates the envelope of the basic signal.

[Aspect 22] The encoding apparatus of Aspect 14, wherein the factor estimator comprises: a first tonality calculating unit to calculate a tonality of a high-frequency section of the frequency domain input signal; a second tonality calculating unit to calculate a tonality of the basic signal; and a factor calculating unit to calculate the energy control factor using the tonality of the high-frequency section and the tonality of the basic signal.

[Aspect 23] The encoding apparatus of Aspect 15, wherein the energy controller controls the extracted energy when the energy control factor is less than a predetermined energy control factor.

[Aspect 24] The encoding apparatus of Aspect 15, wherein the energy quantizer selects a sub-vector of an energy vector, quantizes the selected sub-vector, and quantizes non-selected sub-vectors using an interpolation error.

[Aspect 25] The encoding apparatus of Aspect 24, wherein the energy quantizer selects the sub-vector and quantizes the selected sub-vector at regular intervals.

[Aspect 26] The encoding apparatus of Aspect 24, wherein the energy quantizer selects candidates for the sub-vector, and performs multi-stage Vector Quantization (VQ) including at least two stages.

[Aspect 27] An encoding apparatus, comprising: a processor to control one or more processor-executable units; an energy extractor to extract an energy from a frequency domain input signal, based on an encoding mode; an energy controller to control the extracted energy based on the encoding mode; and an energy quantizer to quantize the controlled energy based on the encoding mode.

[Aspect 28] An encoding apparatus, comprising: a processor to control one or more processor-executable units; an encoding mode selecting unit to select an encoding mode of bandwidth extension encoding using a frequency domain input signal and a time domain input signal; and an extension encoding unit to perform the bandwidth extension encoding using the frequency domain input signal and the selected encoding mode.

[Aspect 29] The encoding apparatus of Aspect 28, wherein the encoding mode selecting unit classifies the frequency domain input signal using the frequency domain input signal and the time domain input signal, determines the encoding mode of the bandwidth extension encoding mode based on information obtained by classifying the frequency domain input signal, and determines a number of frequency bands based on the determined encoding mode.

[Aspect 30] The encoding apparatus of Aspect 28, wherein the extension encoding unit comprises: an energy extractor to extract an energy from the frequency domain input signal, based on the encoding mode; an energy controller to control the extracted energy based on the encoding mode; and an energy quantizer to quantize the controlled energy based on the encoding mode.

[Aspect 31] The encoding apparatus of Aspect 30, wherein the energy extractor extracts an energy corresponding to a frequency band based on the encoding mode.

[Aspect 32] The encoding apparatus of Aspect 30, wherein the energy controller controls the energy using an energy control factor estimated based on a basic signal of the frequency domain input signal.

[Aspect 33] The encoding apparatus of Aspect 30, wherein the energy quantizer quantizes the energy using a scheme optimized for the frequency domain input signal, based on the encoding mode.

[Aspect 34] The encoding apparatus of Aspect 33, wherein the energy quantizer quantizes the energy by applying a frequency weighting method to a frequency band, when the encoding mode is a transient mode.



[Aspect 35] The encoding apparatus of Aspect 34, wherein the frequency weighting method is used to assign a weight to a low-frequency band having a high perceptual importance, and to quantize the energy.

[Aspect 36] The encoding apparatus of Aspect 33, wherein the energy quantizer quantizes the energy by applying an unequal bit allocation method to a frequency band, when the encoding mode is a normal mode or a harmonic mode.

[Aspect 37] The encoding apparatus of Aspect 36, wherein the unequal bit allocation method is used to allocate a greater number of bits to a low-frequency band having a high perceptual importance, than to a high-frequency band, and to quantize the energy.

[Aspect 38] The encoding apparatus of Aspect 30, wherein the energy quantizer predicts a representative value of a quantization target vector including at least two elements, and performs Vector Quantization (VQ) on an error signal between the predicted representative value and each of the at least two elements of the quantization target vector.

[Aspect 39] A decoding apparatus, comprising: a processor to control one or more processor-executable units; a core-decoding unit to core-decode a time domain input signal included in a bitstream; an up-sampling unit to up-sample the core-decoded time domain input signal; a frequency transforming unit to transform the up-sampled time domain input signal to a frequency domain input signal; and an extension decoding unit to perform bandwidth extension decoding, using an energy of the frequency domain input signal.

[Aspect 40] The decoding apparatus of Aspect 39, wherein the extension decoding unit comprises: an inverse-quantizer to inverse-quantize the energy of the time domain input signal; a basic signal generator to generate a basic signal using the frequency domain input signal; a gain calculating unit to calculate a gain using the inverse-quantized energy and an energy of the basic signal, the gain being applied to the basic signal; and a gain applier to apply the calculated gain for each frequency band.

[Aspect 41] The decoding apparatus of Aspect 40, wherein the inverse-quantizer selects a sub-vector of an energy vector, inverse-quantizes the selected sub-vector, interpolates the inverse-quantized sub-vector, adds an interpolation error value to the interpolated sub-vector, and finally inverse-quantizes the energy.

[Aspect 42] The decoding apparatus of Aspect 40, wherein the basic signal generator comprises: an artificial signal generator to generate an artificial signal corresponding to a high-frequency section by copying and folding a low-frequency section of the frequency domain input signal; an envelope estimator to estimate an envelope of the basic signal using a window contained in the bitstream; and an envelope applier to apply the estimated envelope to the artificial signal.

[Aspect 43] The decoding apparatus of Aspect 40, wherein the gain calculating unit and the gain applier set a sub-band used to apply energy smoothing, and generate an energy for each sub-band using an interpolation, and wherein the gain is calculated for each sub-band.

[Aspect 44] An encoding method, comprising: core-encoding a down-sampled time domain input signal; transforming the time domain input signal to a frequency domain input signal; and performing bandwidth extension encoding, by way of a processor, using a basic signal of the frequency domain input signal.

[Aspect 45] The encoding method of Aspect 44, wherein the performing comprises: generating the basic signal of the frequency domain input signal, using a frequency spectrum of the frequency domain input signal; estimating an energy control factor using the basic signal; extracting an energy from the frequency domain input signal; controlling the extracted energy using the energy control factor; and quantizing the controlled energy.

[Aspect 46] The encoding method of Aspect 45, wherein the generating comprises: generating an artificial signal corresponding to a high-frequency section by copying and folding a low-frequency section of the frequency domain input signal; estimating an envelope of the basic signal using a window; and applying the estimated envelope to the artificial signal.

[Aspect 47] The encoding method of Aspect 46, wherein a peak of the window corresponds to a frequency index where an envelope of the basic signal is to be estimated, and wherein the estimating comprises selecting a window based on a comparison result of either a tonality or a correlation, and estimating the envelope of the basic signal using the selected window.

[Aspect 48] The encoding method of Aspect 46, wherein the estimating comprises estimating an average of frequency magnitudes of a whitening band as an envelope of a frequency contained in the whitening band.

[Aspect 49] The encoding method of Aspect 48, wherein the estimating comprises controlling a number of frequency spectrums contained in the whitening band based on a core-encoding mode, and estimating the envelope of the basic signal.

[Aspect 50] The encoding method of Aspect 45, wherein the estimating comprises: calculating a tonality of a high-frequency section of the frequency domain input signal; calculating a tonality of the basic signal; and calculating the energy control factor using the tonality of the high-frequency section and the tonality of the basic signal.

[Aspect 51] The encoding method of Aspect 45, wherein the controlling comprises controlling the extracted energy when the energy control factor is less than a predetermined energy control factor.

[Aspect 52] The encoding method of Aspect 45, wherein the quantizing comprises selecting a sub-vector of an

energy vector, quantizing the selected sub-vector, and quantizing non-selected sub-vectors using an interpolation error.

[Aspect 53] The encoding method of Aspect 52, wherein the quantizing comprises selecting the sub-vector and quantizing the selected sub-vector at regular intervals.

[Aspect 54] The encoding method of Aspect 52, wherein the quantizing comprises selecting candidates for the sub-vector, and performing multi-stage Vector Quantization (VQ) including at least two stages.

[Aspect 55] The encoding method of Aspect 54, wherein the quantizing comprises generating an index set for minimizing a Mean Square Error (MSE) or a Weighted Mean Square Error (WMSE) in each stage for each of the candidates for the sub-vector, and selecting candidates for a sub-vector having a smallest sum of an MSE or WMSE in all stages.

[Aspect 56] The encoding method of Aspect 54, wherein the quantizing comprises generating an index set for minimizing an MSE or a WMSE in each stage for each of the candidates for the sub-vector, restoring an energy vector through an inverse-quantization operation, and selecting candidates for a sub-vector for minimizing an MSE or WMSE between the restored energy vector and an original energy vector.

[Aspect 57] An encoding method, comprising: core-encoding a down-sampled time domain input signal; transforming the core-encoded time domain input signal to a frequency domain input signal; and performing bandwidth extension encoding, by way of a processor, using characteristics of the frequency domain input signal, and using a basic signal of the frequency domain input signal.

[Aspect 58] The encoding method of Aspect 57, wherein the performing comprises: generating the basic signal of the frequency domain input signal, using a frequency spectrum of the frequency domain input signal; estimating an energy control factor using the basic signal and the characteristics of the frequency domain input signal; extracting an energy from the frequency domain input signal; controlling the extracted energy using the energy control factor; and quantizing the controlled energy.

[Aspect 59] The encoding method of Aspect 58, wherein the performing further comprises classifying the frequency domain input signal based on characteristics of the frequency domain input signal, using the frequency spectrum of the frequency domain input signal, and wherein the estimating comprises estimating the energy control factor using the characteristics of the frequency domain input signal determined in the classifying.

[Aspect 60] The encoding method of Aspect 58, wherein the estimating comprises estimating the energy control factor using the characteristics of the frequency domain input signal determined in the core-encoding.

[Aspect 61] The encoding method of Aspect 58, wherein the generating comprises: generating an artificial signal corresponding to a high-frequency section by copying and folding a low-frequency section of the frequency domain input signal; estimating an envelope of the basic signal using a window; and applying the estimated envelope to the artificial signal.

[Aspect 62] The encoding method of Aspect 61, wherein a peak of the window corresponds to a frequency index where an envelope of the basic signal is to be estimated, and wherein the estimating comprises selecting a window based on a comparison result of either a tonality or a correlation, and estimating the envelope of the basic signal using the selected window.

[Aspect 63] The encoding method of Aspect 61, wherein the estimating comprises estimating an average of frequency magnitudes of a whitening band as an envelope of a frequency contained in the whitening band.

[Aspect 64] The encoding method of Aspect 63, wherein the estimating comprises controlling a number of frequency spectrums contained in the whitening band based on a core-encoding mode, and estimating the envelope of the basic signal.

[Aspect 65] The encoding method of Aspect 58, wherein the estimating comprises: calculating a tonality of a high-frequency section of the frequency domain input signal; calculating a tonality of the basic signal; and calculating the energy control factor using the tonality of the high-frequency section and the tonality of the basic signal.

[Aspect 66] The encoding method of Aspect 58, wherein the controlling comprises controlling the extracted energy when the energy control factor is less than a predetermined energy control factor.

[Aspect 67] The encoding method of Aspect 58, wherein the quantizing comprises selecting a sub-vector of an energy vector, quantizing the selected sub-vector, and quantizing non-selected sub-vectors using an interpolation error.

[Aspect 68] The encoding method of Aspect 67, wherein the quantizing comprises selecting the sub-vector and quantizing the selected sub-vector at regular intervals.

[Aspect 69] The encoding method of Aspect 67, wherein the quantizing comprises selecting candidates for the sub-vector, and performing multi-stage Vector Quantization (VQ) including at least two stages.

[Aspect 70] An encoding method, comprising: extracting an energy from a frequency domain input signal, based on an encoding mode; controlling the extracted energy based on the encoding mode; and quantizing, by way of a processor, the controlled energy based on the encoding mode.

[Aspect 71] An encoding method, comprising: selecting an encoding mode of bandwidth extension encoding using

a frequency domain input signal and a time domain input signal; and performing the bandwidth extension encoding, by way of a processor, using the frequency domain input signal and the selected encoding mode.

[Aspect 72] The encoding method of Aspect 71, wherein the selecting comprises: classifying the frequency domain input signal using the frequency domain input signal and the time domain input signal; and determining the encoding mode of the bandwidth extension encoding mode based on information obtained by classifying the frequency domain input signal, and determining a number of frequency bands based on the determined encoding mode.

[Aspect 73] The encoding method of Aspect 71, wherein the performing comprises: extracting an energy from the frequency domain input signal, based on the encoding mode; controlling the extracted energy based on the encoding mode; and quantizing the controlled energy based on the encoding mode.

[Aspect 74] The encoding method of Aspect 73, wherein the extracting comprises extracting an energy corresponding to a frequency band based on the encoding mode.

[Aspect 75] The encoding method of Aspect 73, wherein the controlling comprises controlling the energy using an energy control factor estimated based on a basic signal of the frequency domain input signal.

[Aspect 76] The encoding method of Aspect 73, wherein the quantizing comprises quantizing the energy using a scheme optimized for the frequency domain input signal, based on the encoding mode.

[Aspect 77] The encoding method of Aspect 76, wherein the quantizing comprises quantizing the energy by applying a frequency weighting method to a frequency band, when the encoding mode is a transient mode.

[Aspect 78] The encoding method of Aspect 77, wherein the frequency weighting method is used to assign a weight to a low-frequency band having a high perceptual importance, and to quantize the energy.

[Aspect 79] The encoding method of Aspect 76, wherein the quantizing comprises quantizing the energy by applying an unequal bit allocation method to a frequency band, when the encoding mode is a normal mode or a harmonic mode.

[Aspect 80] The encoding method of Aspect 79, wherein the unequal bit allocation method is used to allocate a greater number of bits to a low-frequency band having a high perceptual importance, than to a high-frequency band, and to quantize the energy.

[Aspect 81] The encoding method of Aspect 73, wherein the quantizing comprises predicting a representative value of a quantization target vector including at least two elements, and performing Vector Quantization (VQ) on an error signal between the predicted representative value and each of the at least two elements of the quantization target vector.

[Aspect 82] A decoding method, comprising: core-decoding a time domain input signal included in a bitstream; up-sampling the core-decoded time domain input signal; transforming the up-sampled time domain input signal to a frequency domain input signal; and performing bandwidth extension decoding, by way of a processor, using an energy of the time domain input signal and using the frequency domain input signal.

[Aspect 83] The decoding method of Aspect 82, wherein the performing comprises: inverse-quantizing the energy of the time domain input signal; generating a basic signal using the frequency domain input signal; calculating a gain using the inverse-quantized energy and an energy of the basic signal, the gain being applied to the basic signal; and applying the calculated gain for each frequency band.

[Aspect 84] The decoding method of Aspect 83, wherein the inverse-quantizing comprises selecting a sub-vector of an energy vector, inverse-quantizing the selected sub-vector, interpolating the inverse-quantized sub-vector, adding an interpolation error value to the interpolated sub-vector, and finally inverse-quantizing the energy.

[Aspect 85] The decoding method of Aspect 84, wherein the generating comprises: generating an artificial signal corresponding to a high-frequency section by copying and folding a low-frequency section of the frequency domain input signal; estimating an envelope of the basic signal using a window contained in the bitstream; and applying the estimated envelope to the artificial signal.

[Aspect 86] The decoding method of Aspect 84, wherein the calculating comprises setting a sub-band used to apply energy smoothing, and generating an energy for each sub-band through an interpolation, and wherein the gain is calculated for each sub-band.

[Aspect 87] A non-transitory computer readable recording medium storing a program to cause a computer to implement the method of Aspect 44.

## Claims

### 1. An audio encoding method comprising:

core-encoding a down-sampled time domain input signal;  
transforming the time domain input signal to a frequency domain input signal;  
generating a basic signal forming a fine structure of a low-frequency domain,  
using a frequency spectrum of the frequency domain input signal;

obtaining an energy control factor using the basic signal;  
 obtaining an energy from the frequency domain input signal;  
 controlling the obtained energy using the obtained energy control factor; and  
 quantizing the controlled energy,  
 wherein the quantizing of the controlled energy comprises:

selecting a sub-vector of an energy vector of the controlled energy;  
 quantizing and inverse-quantizing the selected sub-vector;  
 interpolating non-selected sub-vectors using the inverse-quantized sub-vector;  
 calculating an interpolation error denoting a difference between the interpolated non-selected sub-vectors  
 and the original non-selected sub-vectors; and  
 quantizing the interpolation error.

2. The audio encoding method of claim 1, wherein the obtaining of the energy control factor is based on a ratio between tonality of the basic signal and tonality of the frequency domain input signal.

3. The audio encoding method of claim 1, wherein the quantizing of the controlled energy comprises quantizing the controlled energy based on a weighted mean square error (WMSE).

4. The audio encoding method of claim 1, wherein the quantizing of the controlled energy comprises quantizing the controlled energy by using a multi-stage vector quantization.

5. The audio encoding method of claim 1, wherein the selecting of the subvector comprises selecting the sub-vector at regular intervals.

6. An audio encoding apparatus comprising at least one processor configured to:

core-encode a down-sampled time domain input signal,  
 transform the time domain input signal to a frequency domain input signal,  
 generate a basic signal forming a fine structure of a low-frequency domain, using a frequency spectrum of the frequency domain input signal,  
 obtain an energy control factor using the basic signal;  
 obtain an energy from the frequency domain input signal,  
 control the obtained energy using the obtained energy control factor, and  
 quantize the controlled energy,  
 wherein the at least one processor configured to quantize the controlled energy by:

selecting a sub-vector of an energy vector of the controlled energy;  
 quantizing and inverse-quantizing the selected sub-vector;  
 interpolating non-selected sub-vectors using the inverse-quantized sub-vector;  
 calculating an interpolation error denoting a difference between the interpolated non-selected sub-vectors  
 and the original non-selected sub-vectors; and  
 quantizing the interpolation error.

7. The audio encoding apparatus of claim 6, wherein the at least one processor configured to obtain the energy control factor based on a ratio between tonality of the basic signal and tonality of the frequency domain input signal.

8. The audio encoding apparatus of claim 6, wherein the at least one processor configured to quantize the controlled energy based on a weighted mean square error (WMSE).

9. The audio encoding apparatus of claim 6, wherein the at least one processor configured to quantize the controlled energy by using a multi-stage vector quantization.

10. The audio encoding apparatus of claim 6, wherein the at least one processor configured to select the sub-vector at regular intervals.

11. A non-transitory computer readable recording medium storing a program to cause a computer adapted to implement the method of any one of claims 1 to 5.

FIG. 1

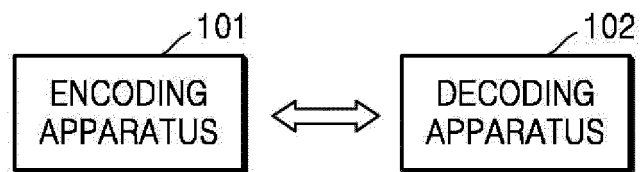


FIG. 2

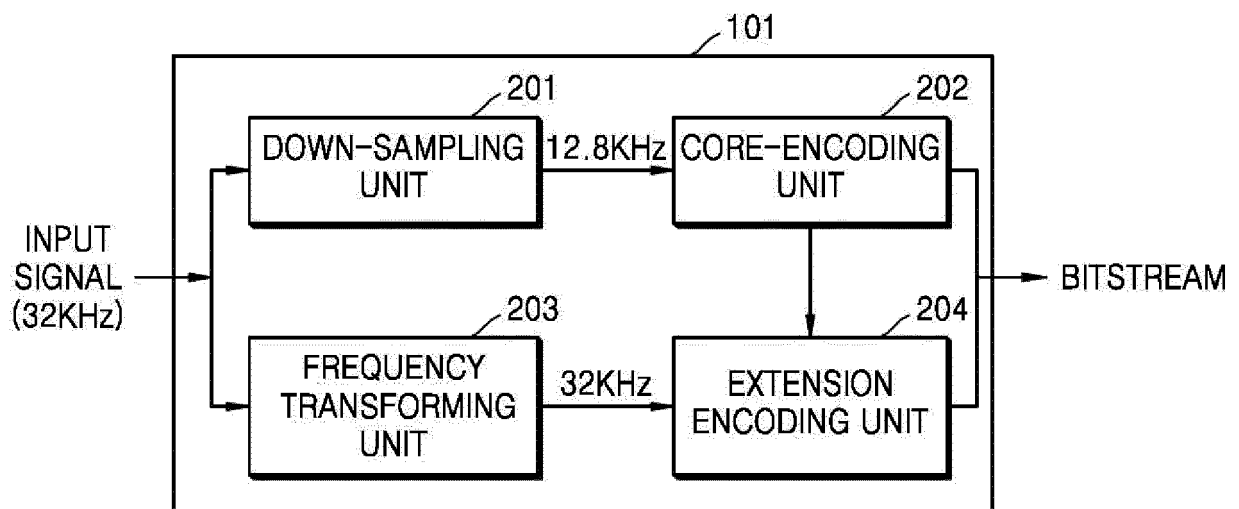


FIG. 3

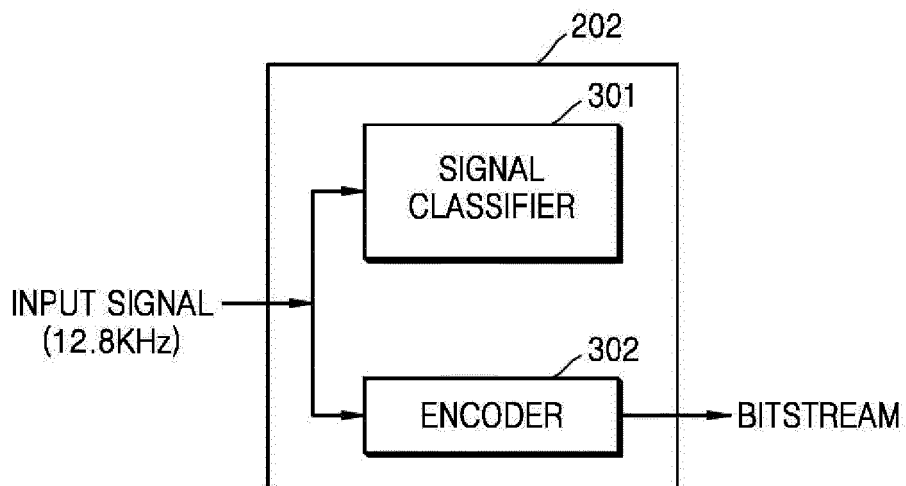


FIG. 4

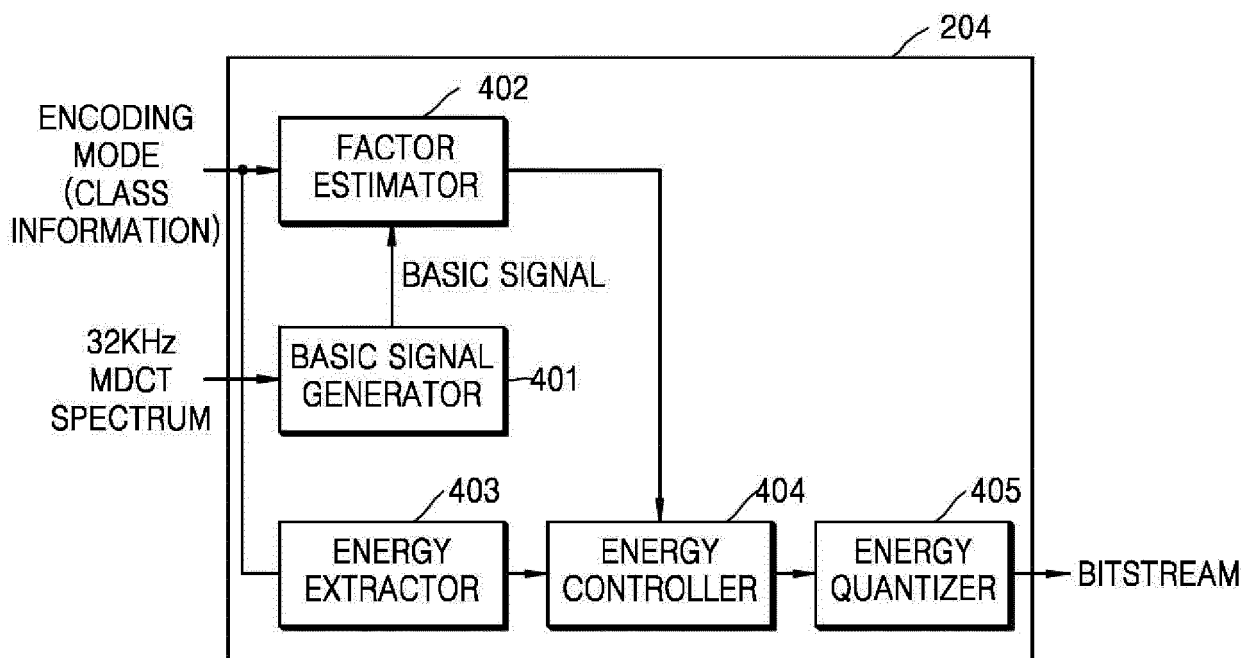


FIG. 5

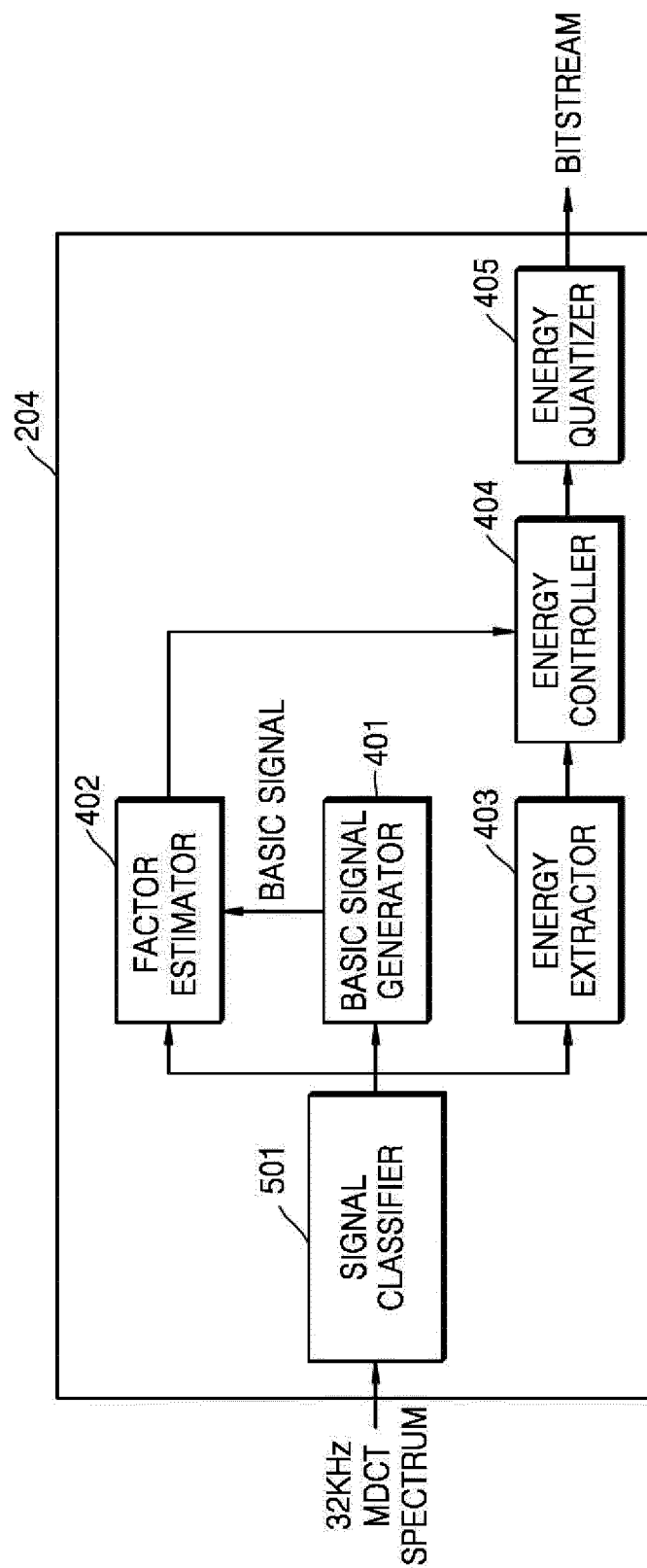


FIG. 6

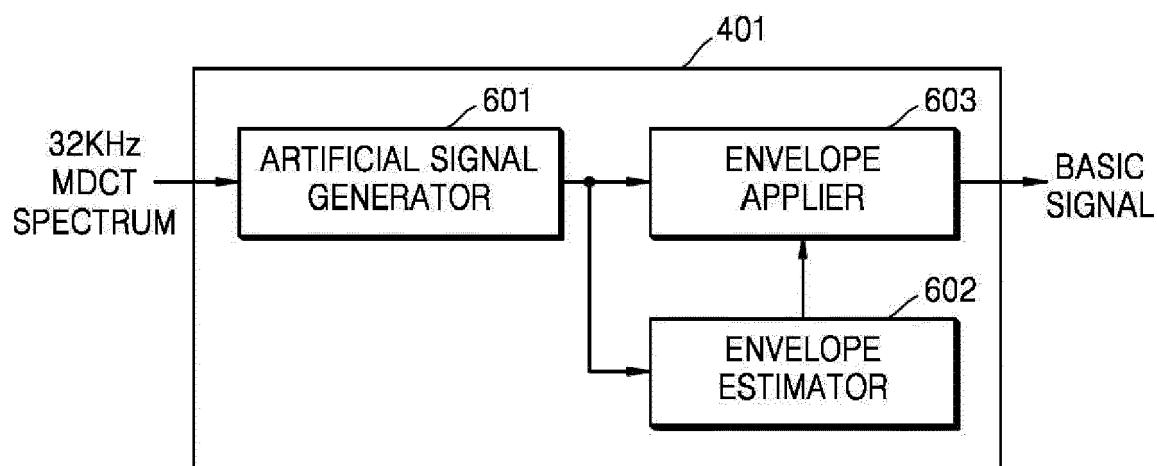


FIG. 7

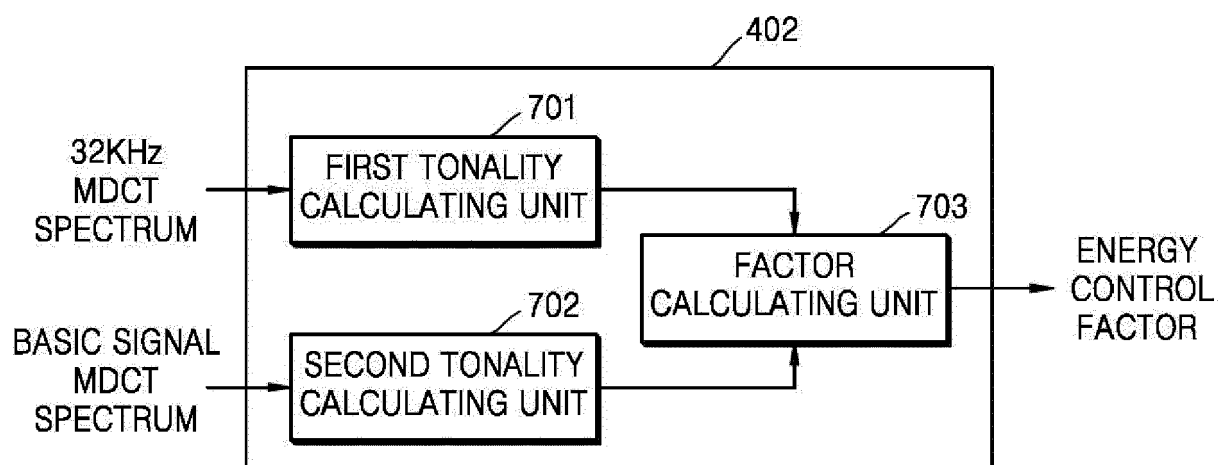




FIG. 8

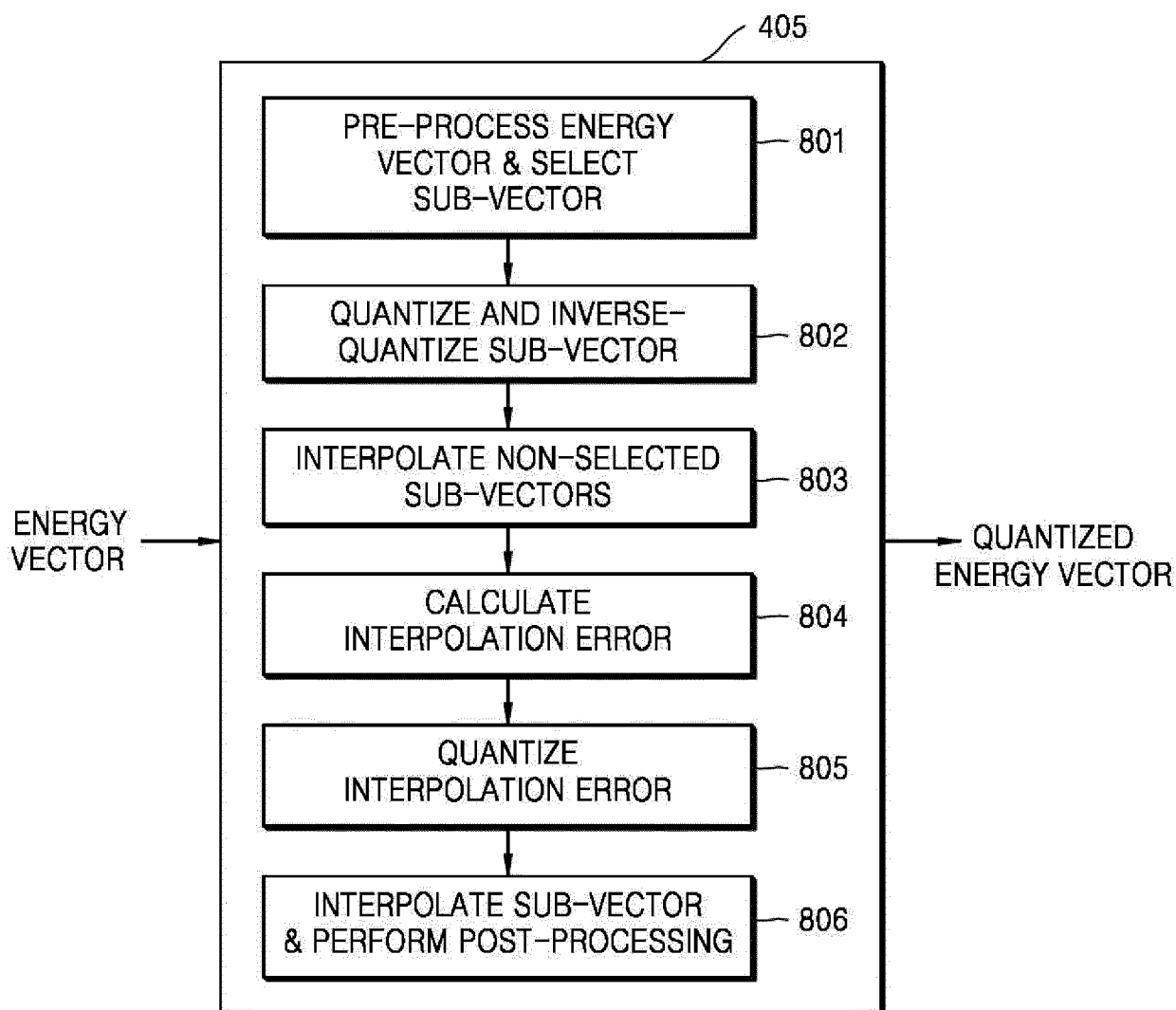


FIG. 9

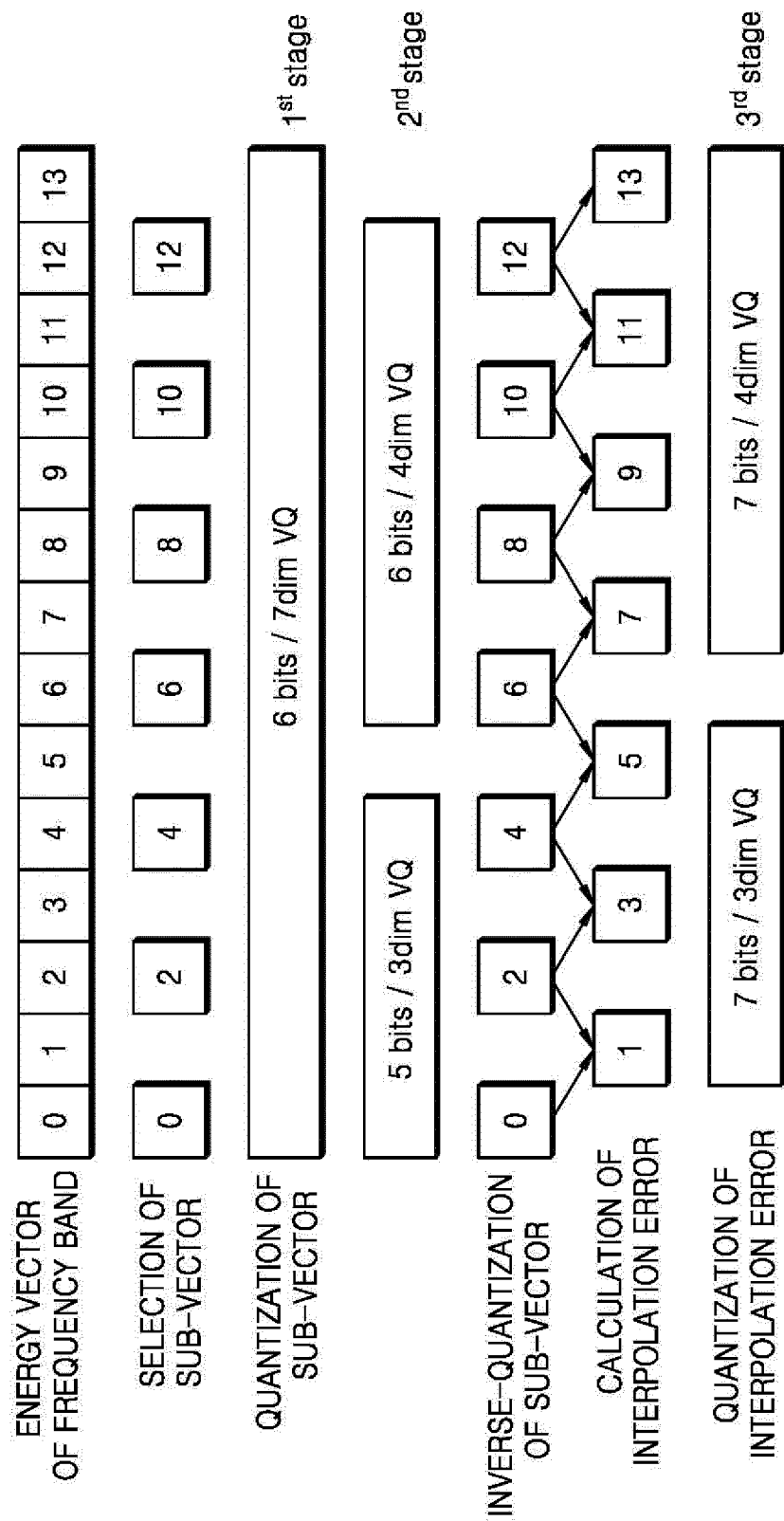


FIG. 10

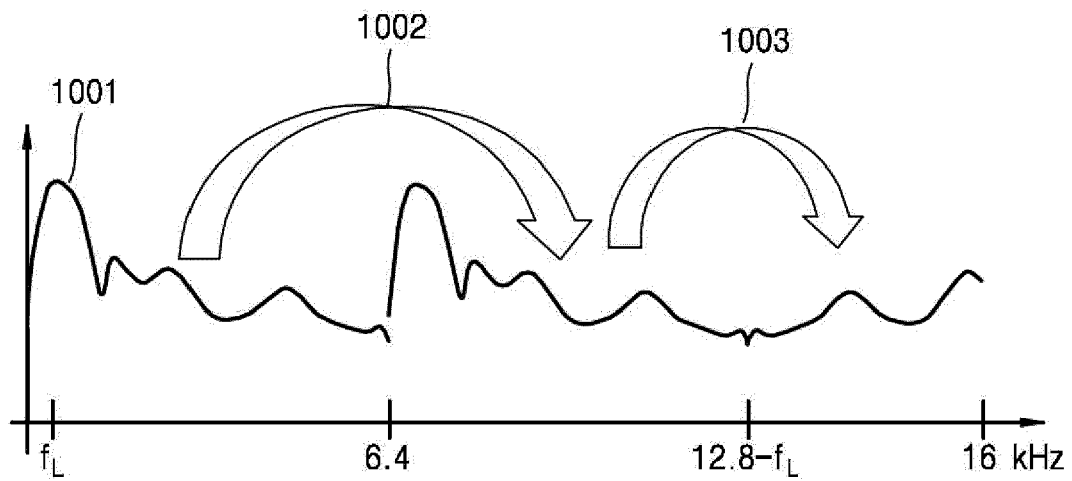


FIG. 11A

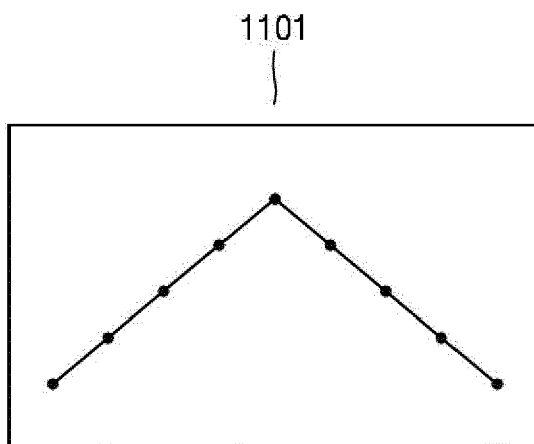


FIG. 11B

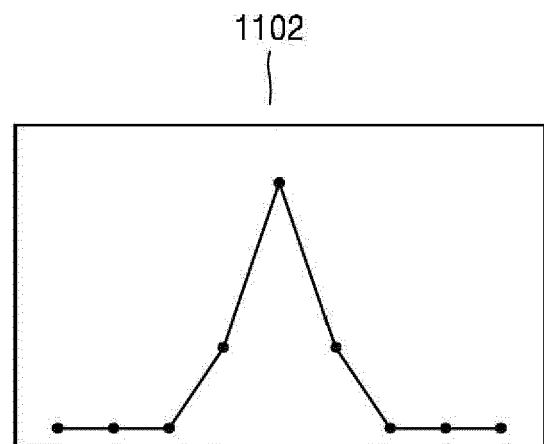


FIG. 12

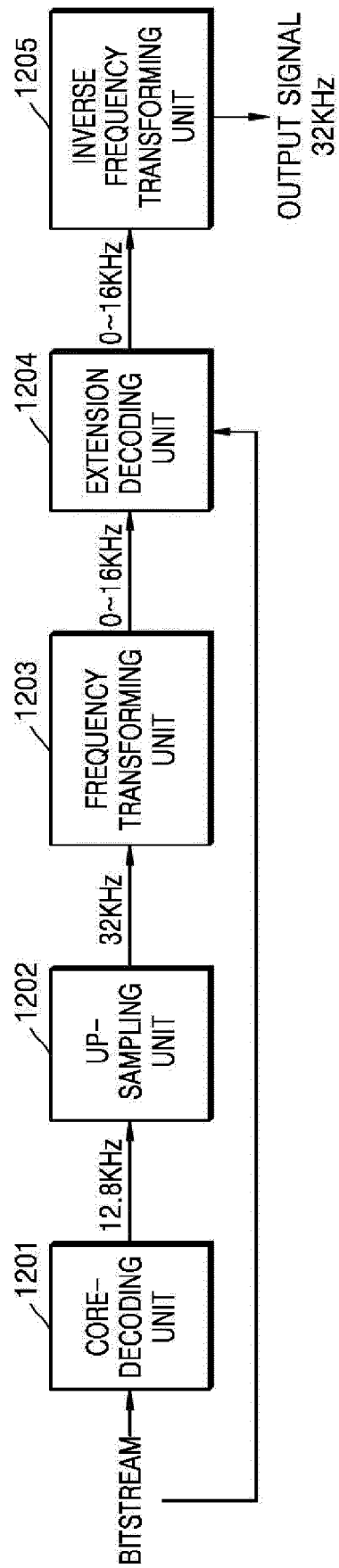


FIG. 13

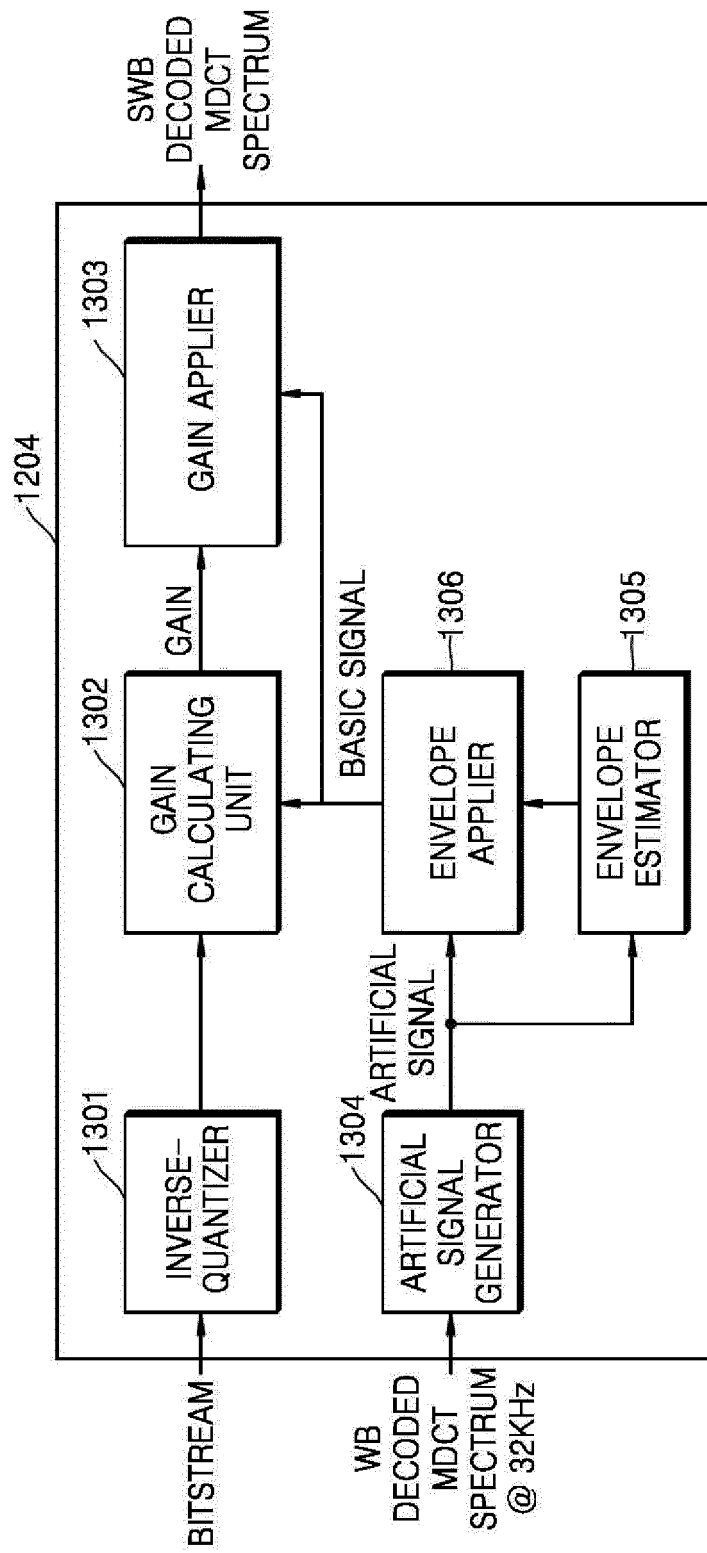


FIG. 14

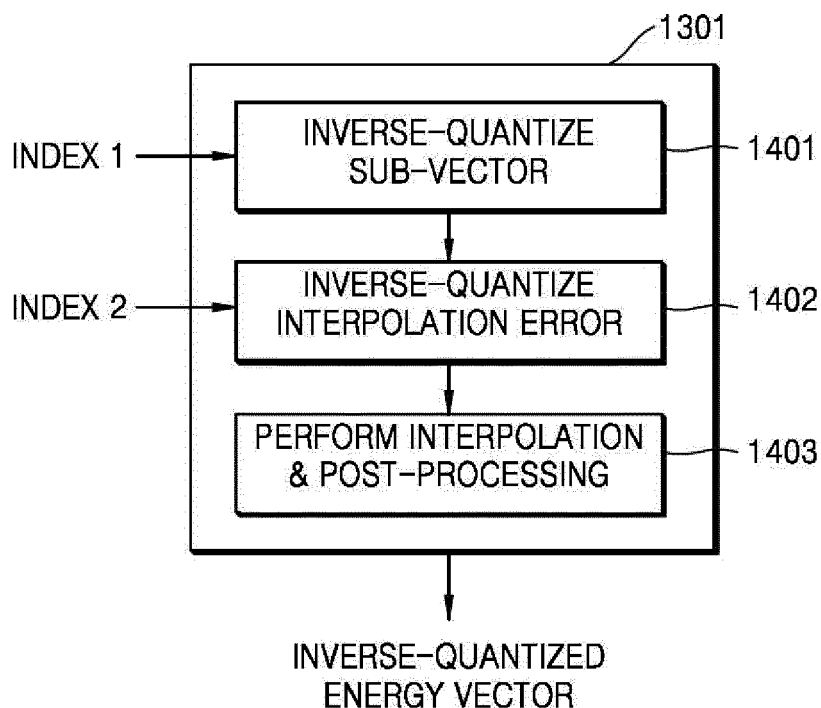


FIG. 15

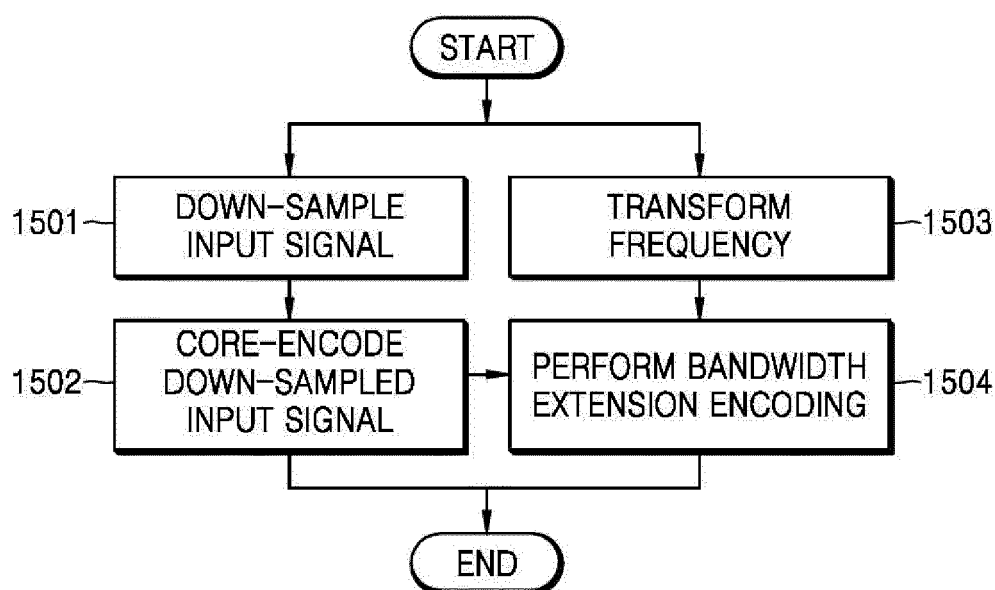


FIG. 16

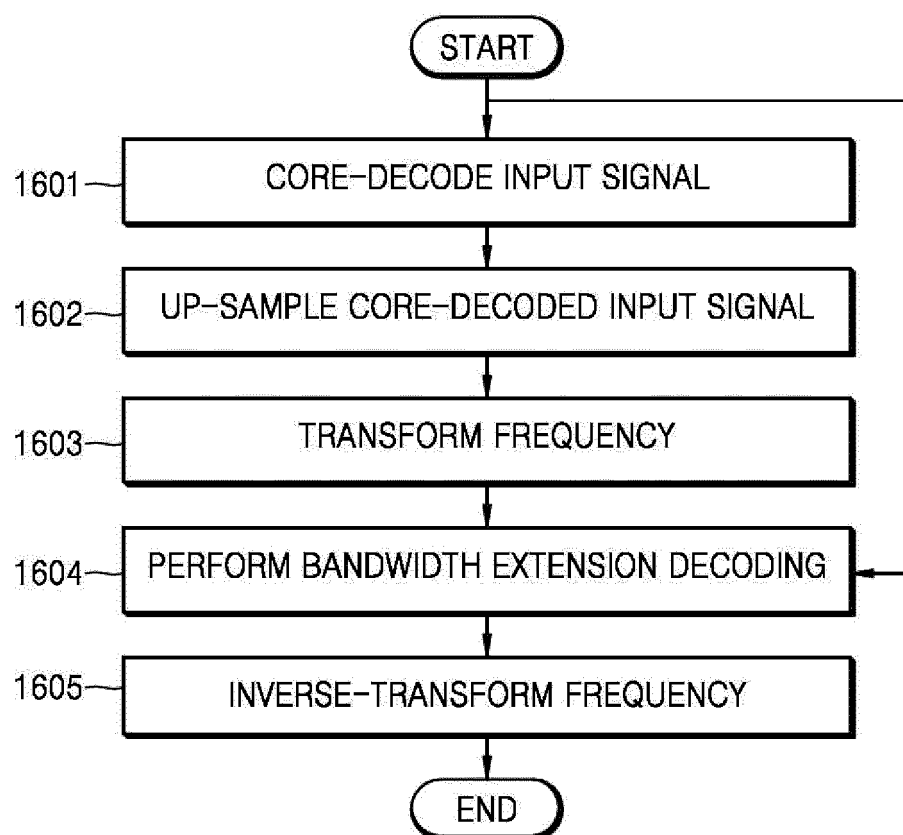


FIG. 17

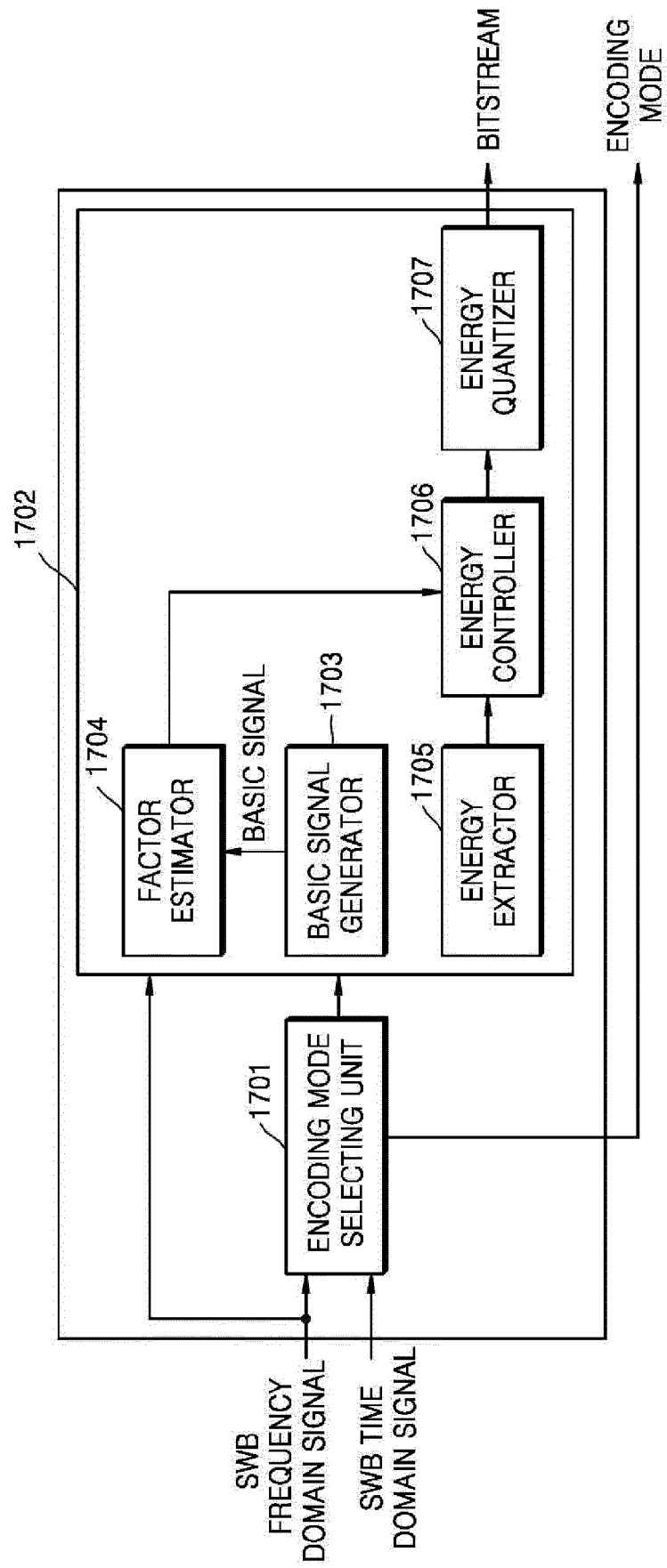




FIG. 18

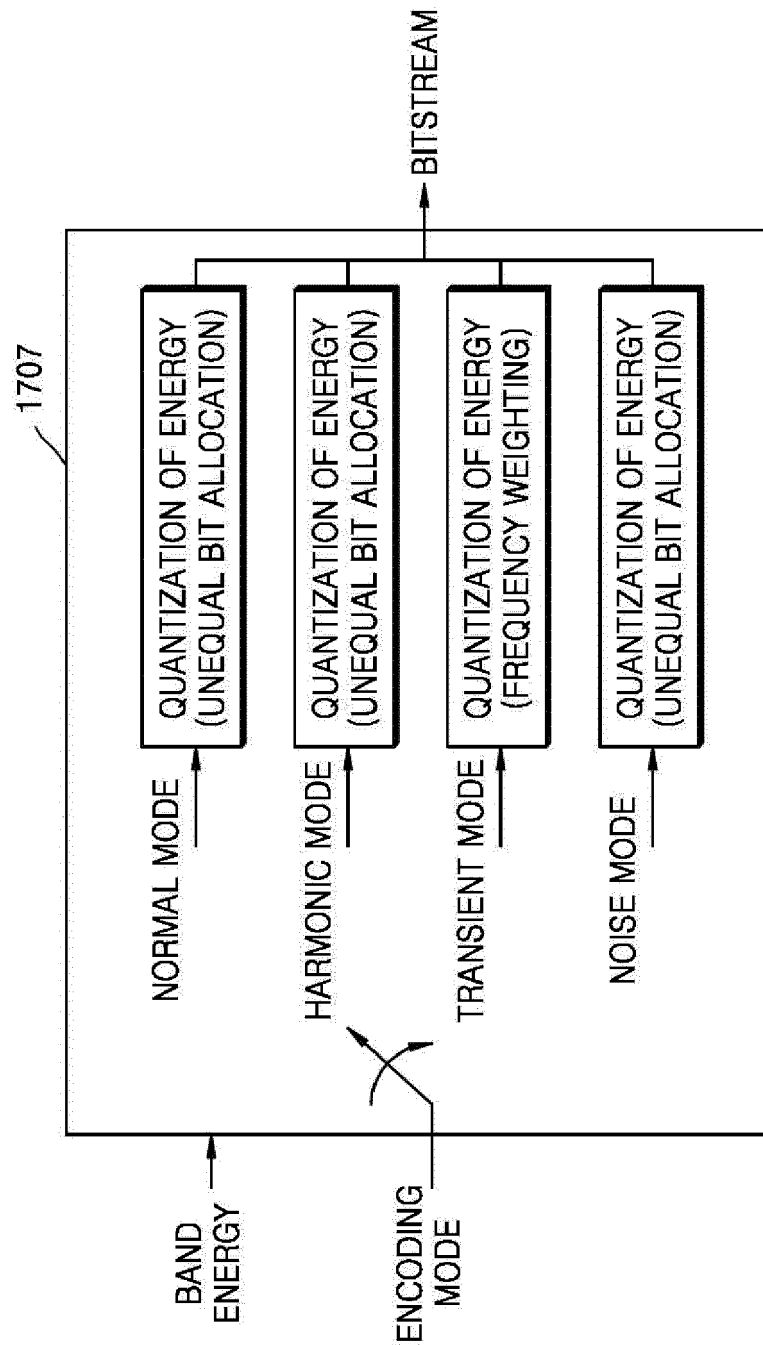


FIG. 19

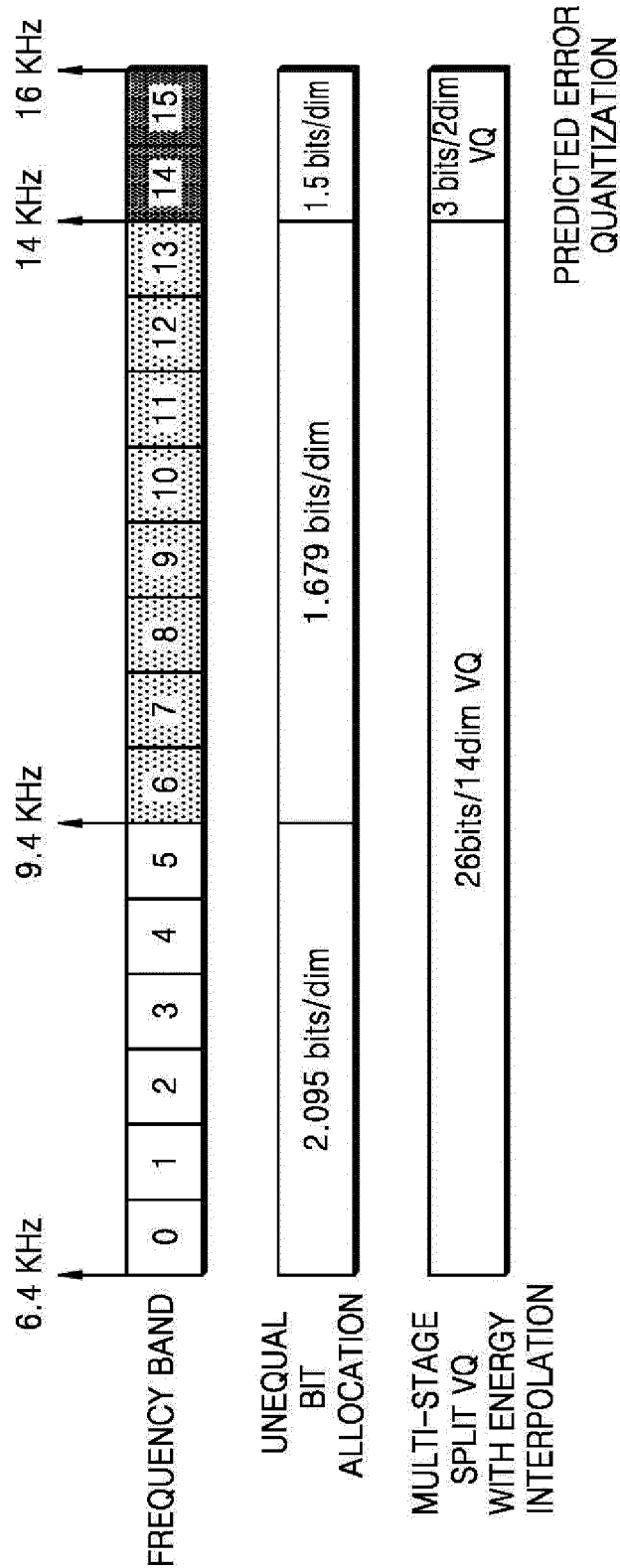


FIG. 20

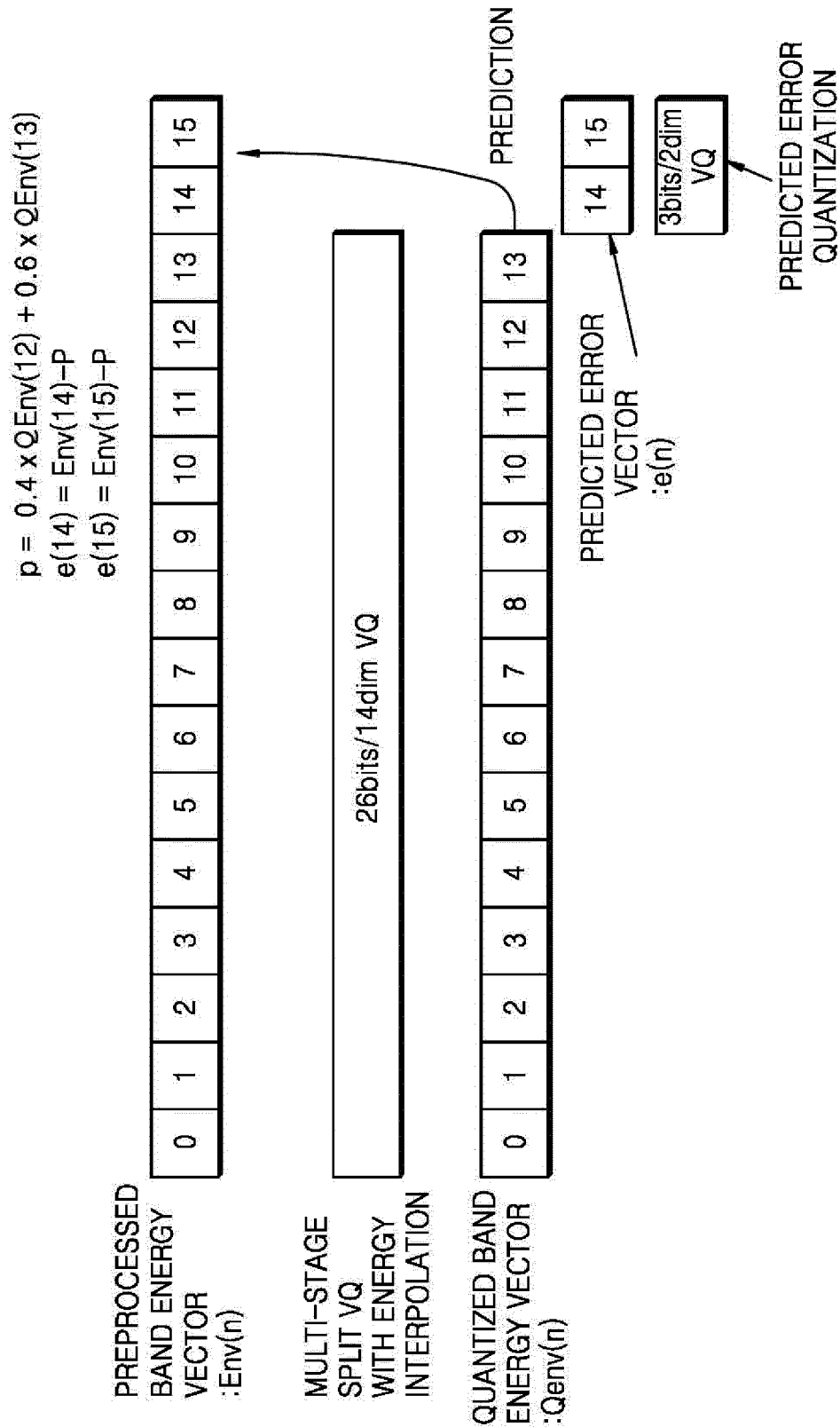


FIG. 21

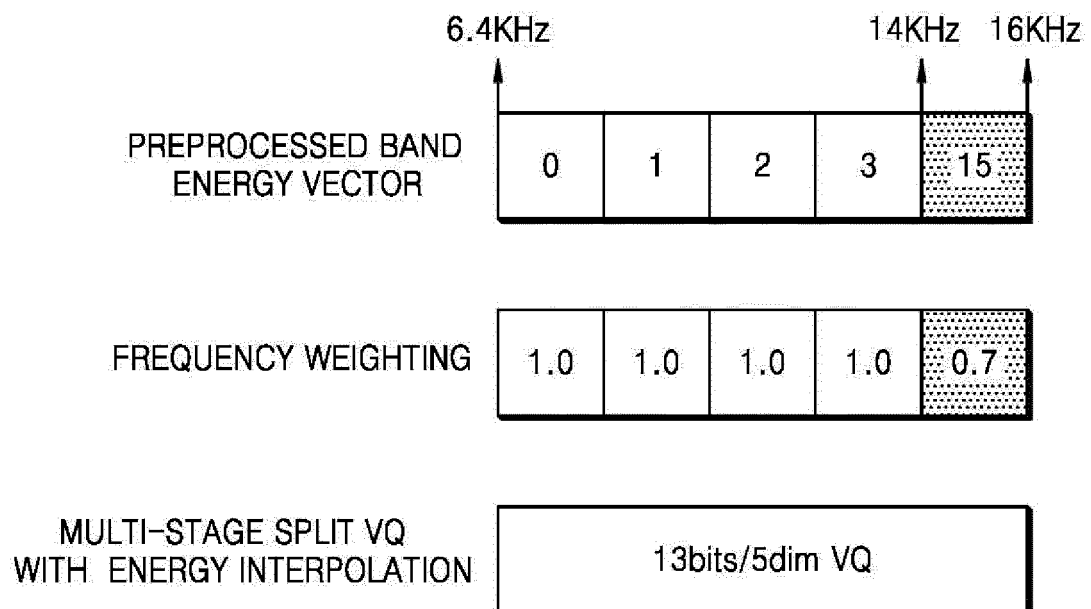


FIG. 22

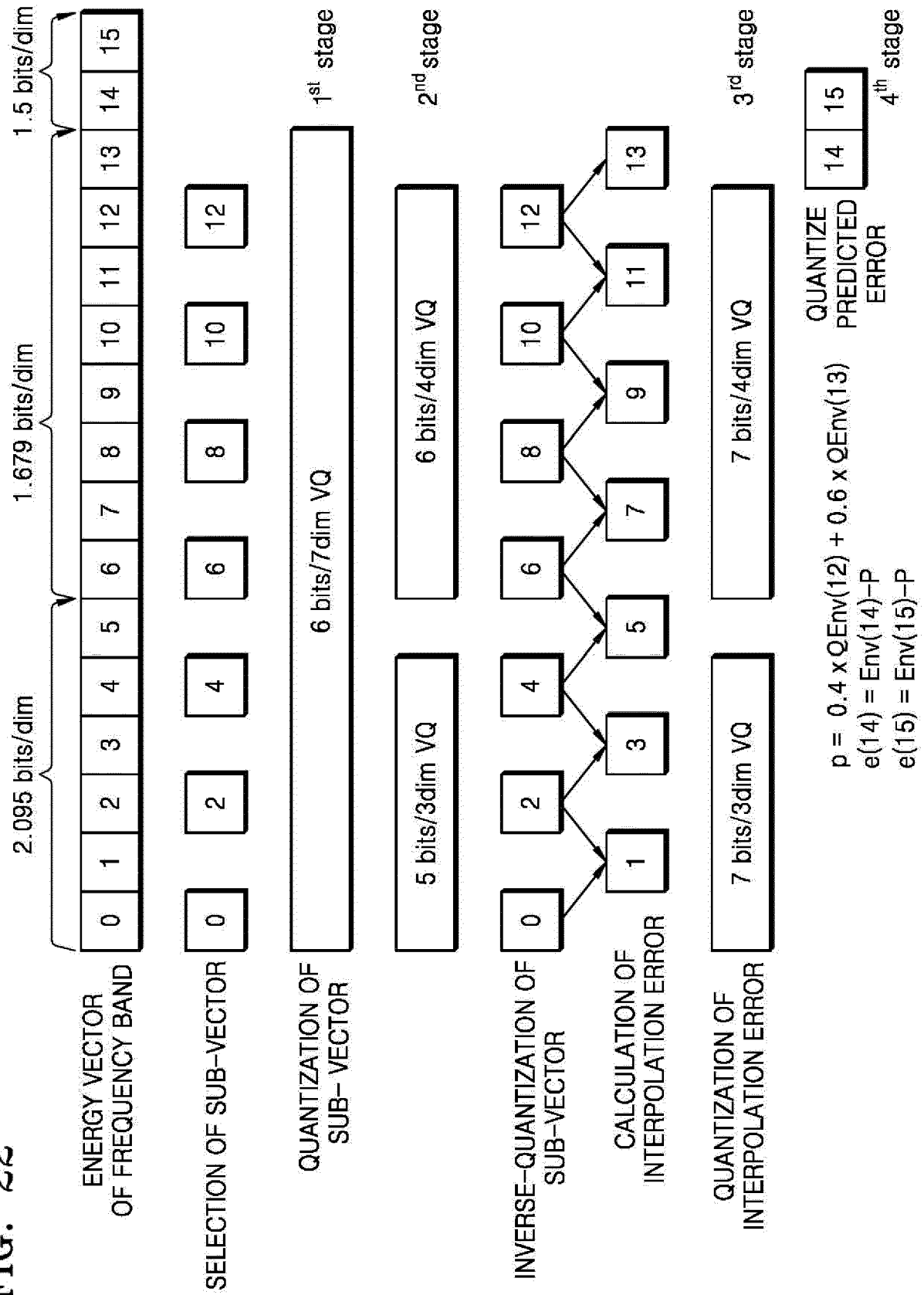


FIG. 23

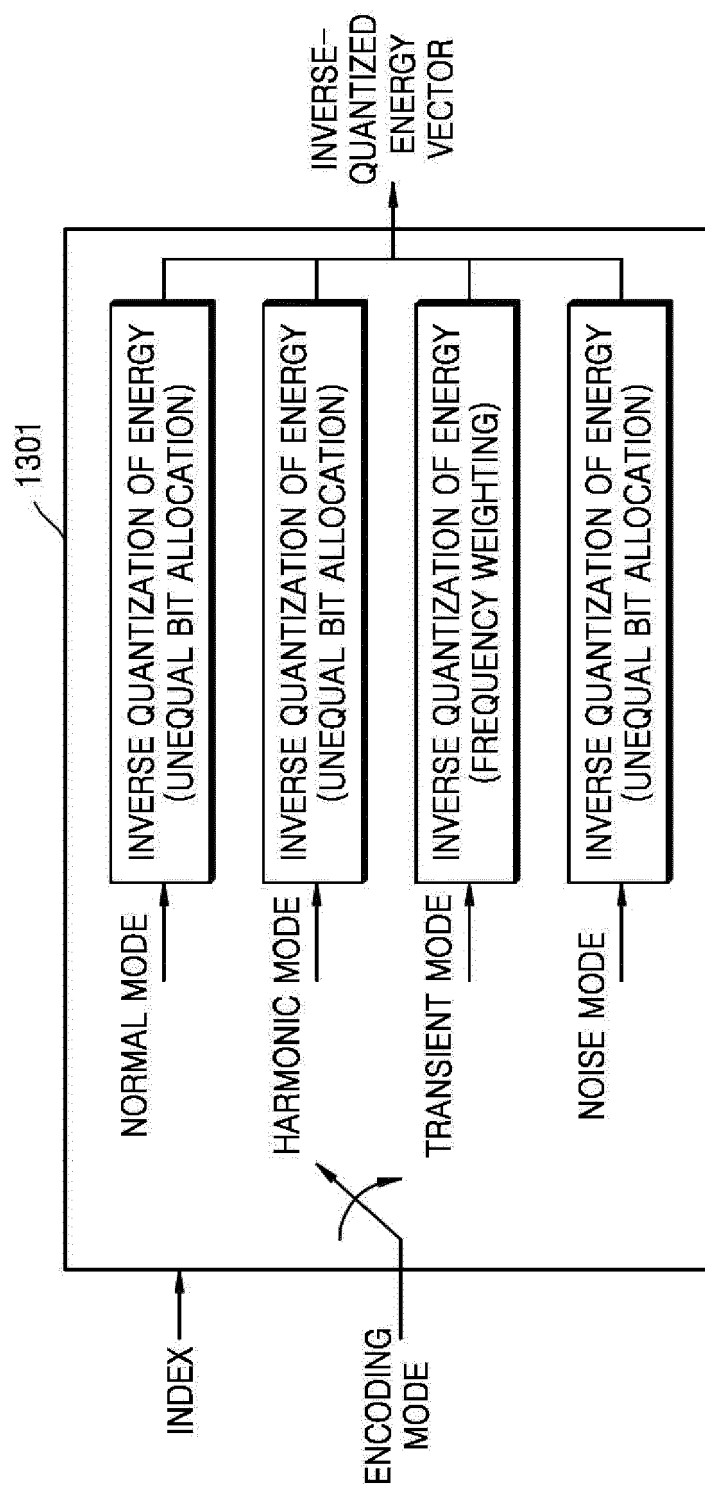
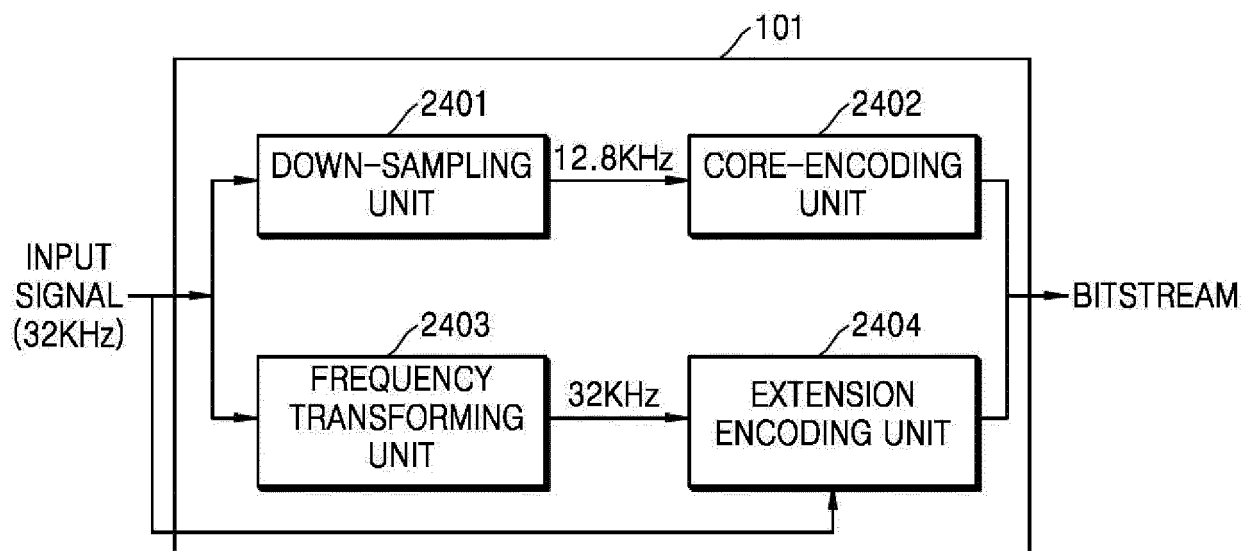


FIG. 24





## EUROPEAN SEARCH REPORT

Application Number  
EP 20 18 5468

5

10

15

20

25

30

35

40

45

50

55

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
A	OSHIKIRI M ET AL: "Efficient spectrum coding for super-wideband speech and its application to 7/10/15 KHz bandwidth scalable coders", ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, 2004. PROCEEDINGS. (ICASSP ' 04). IEEE INTERNATIONAL CONFERENCE ON MONTREAL, QUEBEC, CANADA 17-21 MAY 2004, PISCATAWAY, NJ, USA, IEEE, PISCATAWAY, NJ, USA, vol. 1, 17 May 2004 (2004-05-17), pages 481-484, XP010717670, DOI: 10.1109/ICASSP.2004.1326027 ISBN: 978-0-7803-8484-2 * abstract; figures 1-4 * * page 481, left-hand column, line 28 - page 483, right-hand column, line 10 *	1-11	INV. G10L19/24 G10L21/038
A	WO 2009/066959 A1 (LG ELECTRONICS INC [KR]; OH HYUN O [KR]; JUNG YANG WON [KR]) 28 May 2009 (2009-05-28) * abstract; figure 1 *	1-11	TECHNICAL FIELDS SEARCHED (IPC)
A	US 6 377 915 B1 (SASAKI SEISHI [JP]) 23 April 2002 (2002-04-23) * abstract; figure 1 *	1-11	G10L
A	WO 2008/084924 A1 (SAMSUNG ELECTRONICS CO LTD [KR]) 17 July 2008 (2008-07-17) * abstract; figure 1 *	1-11	
A	US 2009/234645 A1 (BRUHN STEFAN [SE]) 17 September 2009 (2009-09-17) * abstract; figure 5a *	1-11	
The present search report has been drawn up for all claims			
Place of search Munich		Date of completion of the search 16 October 2020	Examiner Zimmermann, Elko
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document	

EPO FORM 1503 03.82 (P04C01)



**ANNEX TO THE EUROPEAN SEARCH REPORT  
ON EUROPEAN PATENT APPLICATION NO.**

EP 20 18 5468

5

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.  
The members are as contained in the European Patent Office EDP file on  
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

16-10-2020

10

15

20

25

30

35

40

45

50

55

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2009066959 A1	28-05-2009	AU 2008326956 A1	28-05-2009
		AU 2008326957 A1	28-05-2009
		BR P10820488 A2	23-05-2017
		CA 2697830 A1	28-05-2009
		CA 2705968 A1	28-05-2009
		CN 101836250 A	15-09-2010
		CN 101868821 A	20-10-2010
		EP 2210253 A1	28-07-2010
		EP 2218068 A1	18-08-2010
		JP 2011504249 A	03-02-2011
		JP 2011504250 A	03-02-2011
		KR 20100085018 A	28-07-2010
		KR 20100095585 A	31-08-2010
		RU 2010125221 A	27-12-2011
		RU 2010125251 A	27-12-2011
		US 2010211400 A1	19-08-2010
		US 2010274557 A1	28-10-2010
		US 2010305956 A1	02-12-2010
US 6377915 B1	23-04-2002	EP 1037197 A2	20-09-2000
		US 6377915 B1	23-04-2002
WO 2008084924 A1	17-07-2008	CN 101236745 A	06-08-2008
		CN 102637434 A	15-08-2012
		CN 102708873 A	03-10-2012
		EP 2105020 A1	30-09-2009
		JP 5558829 B2	23-07-2014
		JP 5600142 B2	01-10-2014
		JP 5869537 B2	24-02-2016
		JP 6208725 B2	04-10-2017
		JP 2010515946 A	13-05-2010
		JP 2012212170 A	01-11-2012
		JP 2013232018 A	14-11-2013
		JP 2015232733 A	24-12-2015
		KR 20080066473 A	16-07-2008
		US 2008172223 A1	17-07-2008
		US 2010010809 A1	14-01-2010
		US 2012316887 A1	13-12-2012
		WO 2008084924 A1	17-07-2008
US 2009234645 A1	17-09-2009	AT 463028 T	15-04-2010
		CN 101512639 A	19-08-2009
		EP 2062255 A1	27-05-2009
		ES 2343862 T3	11-08-2010

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

